COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 42 (2018) 129–167

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/cogs.12482

Problem-Solving Phase Transitions During Team Collaboration

Travis J. Wiltshire,[a,b] Jonathan E. Butner,[a] Stephen M. Fiore[c]

[a]*Department of Psychology, University of Utah*
[b]*Department of Language and Communication, Centre for Human Interactivity, University of Southern Denmark*
[c]*Department of Philosophy and Institute for Simulation and Training, University of Central Florida*

Received 28 April 2016; received in revised form 7 December 2016; accepted 12 December 2016

Abstract

Multiple theories of problem-solving hypothesize that there are distinct qualitative phases exhibited during effective problem-solving. However, limited research has attempted to identify when transitions between phases occur. We integrate theory on collaborative problem-solving (CPS) with dynamical systems theory suggesting that when a system is undergoing a phase transition it should exhibit a peak in entropy and that entropy levels should also relate to team performance. Communications from 40 teams that collaborated on a complex problem were coded for occurrence of problem-solving processes. We applied a sliding window entropy technique to each team's communications and specified criteria for (a) identifying data points that qualify as peaks and (b) determining which peaks were robust. We used multilevel modeling, and provide a qualitative example, to evaluate whether phases exhibit distinct distributions of communication processes. We also tested whether there was a relationship between entropy values at transition points and CPS performance. We found that a proportion of entropy peaks was robust and that the relative occurrence of communication codes varied significantly across phases. Peaks in entropy thus corresponded to qualitative shifts in teams' CPS communications, providing empirical evidence that teams exhibit phase transitions during CPS. Also, lower average levels of entropy at the phase transition points predicted better CPS performance. We specify future directions to improve understanding of phase transitions during CPS, and collaborative cognition, more broadly.

Keywords: Problem-solving; Collaboration; Team cognition; Dynamical systems; Communication

Correspondence should be sent to Travis J. Wiltshire, Department of Language and Communication, University of Southern Denmark, Campusvej 55, DK – 5230 Odense M, Denmark. E-mail: wiltshire@sdu.dk

1. Introduction

Given the increasing complexity of problems faced by teams of scientists and domain practitioners in modern work environments, there is an ever-present need to gain a comprehensive understanding of complex collaborative problem-solving (CPS; OECD, 2015). The study of problem-solving has traditionally focused solely on individual problem solvers' cognitive capacities, in narrow problem contexts (cf. Newell & Simon, 1972). Over the past decade advances have been made toward understanding how humans solve complex problems (e.g., Fischer, Greiff, & Funke, 2012; Quesada, Kintsch, & Gomez, 2005) and how they do so collaboratively with other people in teams (Fiore, Rosen, Salas, Burke, & Jentsch, 2008; Fiore, Smith-Jentsch, Salas, Warner, & Letsky, 2010; Fiore, Rosen, et al., 2010) and in close relationships (Berg, Johnson, Meegan, & Strough, 2003). The empirical examination of CPS has a long history in psychology (e.g., Steiner, 1974). But much of that work has been on simple tasks such as puzzle problems or brainstorming (e.g., Hill, 1982). For complex CPS, more of the work has been theoretical, as a means of scaling up to real-world problems with a greater number of problem solvers, although there has been some empirical work (e.g., de Montjoye, Stopczynski, Shmueli, Pentland, & Lehmann, 2014; Fiore, Wiltshire, Oglesby, O'Keefe, & Salas, 2014; Strough, Patrick, Swenson, Cheng, & Barnes, 2003).

Despite this variability, consistent within CPS research is the notion that there are phases to problem-solving that are iterated through to complete the problem-solving process (Bales & Strodtbeck, 1951; Fiore, Rosen, et al., 2010; Fisher, 1970; Hirokawa, 1983). However, detecting when transitions from one problem-solving phase to another occur has received little attention. Nonetheless, it remains a critical aspect of CPS (cf. Bales & Strodtbeck, 1951) as it has both theoretical and practical implications.

To redress this gap, we build on prior work examining complex CPS by, first, reviewing our theoretical foundation. This includes detailing the *interactive team cognition* approach (Cooke, Gorman, Myers, & Duran, 2013) and the *macrocognition in teams* model of CPS (Fiore, Rosen, et al., 2010; Fiore, Smith-Jentsch, et al., 2010). We then briefly review empirical work that has been conducted in this tradition. The novel contribution of the work presented here is that, instead of focusing on either the identification of CPS processes (e.g., Fiore et al., 2014; Hutchins & Kendall, 2010), or the degree to which certain frequencies and sequences of CPS processes predict performance (e.g., Rosen, 2010), we provide the first empirical effort aimed at identifying the *transition points* between CPS phases. Doing so requires integration of theory on CPS phases with phase transitions in the study of dynamical systems. Thus, we describe the theory and motivation for our dynamical systems approach to CPS and then detail our method for examining phase transitions during team collaboration.

1.1. Interactive team cognition and collaborative problem-solving

With regard to CPS in teams, studying cognition changes from examination of the internal cognitive process that might be occurring during problem-solving to how team interaction unfolds in dynamic and contextually rich settings. Our approach is aligned with the *interactive team cognition* (Cooke et al., 2013) and *team coordination dynamics* approaches (Gorman, Amazeen, & Cooke, 2010). While traditional studies of team cognition have relegated "team cognition" to the static and accumulated knowledge structures that team members share (e.g., shared mental models; Salas & Fiore, 2004), the interactive team cognition approach posits that the activity of the team and thus the interaction between team members *is* team cognition (Cooke, Gorman, & Rowe, 2004; Gorman, Amazeen, et al., 2010; Gorman, Cooke, & Amazeen, 2010), and that this is observable in the dynamics of teams' interactive behaviors (see also related work on group cognition; e.g., Palermos, 2016; Theiner, Allen, & Goldstone, 2010). Generally, work in this area is essential because it provides new ways to assess team cognition that has been shown to be associated with task performance (Cooke, Gorman, Duran, & Taylor, 2007; Stevens, Galloway, Wang, & Berka, 2012) and that are not subject to issues with self-report team cognition measures (see, e.g., Wildman, Salas, & Scott, 2014). Furthermore, it has been argued that team collaboration necessitates interaction and coordination between social entities (Bedwell et al., 2012) and therefore it must be studied as such. Also, it has been noted that the lack of attempts to "detect qualitative changes ("phase transitions") in the quantity and information content of [team] communication patterns" is a gap in this research area (p. 513; Gorman, Cooke, Amazeen, & Fouse, 2012). Building on this line of theorizing, then, we focus on teams' communicative interaction to identify phase transitions.

1.1.1. Macrocognition in teams

The macrocognitive approach to cognition (Hollnagel, 2002) arose from studies of cognition as it occurs in naturalistic settings as a reaction to the limited utility some researchers found when trying to translate laboratory findings of cognition to real-world applications (Klein et al., 2003). On this view, cognition is generally conceptualized as distributed, contextually embedded within a social environment, and dynamic (Hollnagel, 2002; see also Klein et al., 2003) and it is closely related to the distributed cognition approach (Hutchins, 1995). The interactive team cognition approach is also consistent with this view, although it focuses specifically on the dynamics of team members' behavior as indicative of *team* cognition.

The goal of the macrocognition in teams model is to characterize how teams build knowledge during CPS and reach effective problem-solving outcomes via integration of member inputs. In this theorizing, CPS involves the coordination of actions among individuals as they adapt their existing knowledge or generate new knowledge to solve novel and complex problems (Fiore, Smith-Jentsch, et al., 2010; Fiore et al., 2014; Roschelle & Teasley, 1995; Wiltshire, Rosch, Fiorella, & Fiore, 2014).

The macrocognition in teams model is an interdisciplinary integration of prior theorizing on team cognition and group problem-solving (see Fiore, Rosen, et al., 2010; Fiore, Smith-Jentsch, et al., 2010). It consists of five major components that characterize the CPS process: individual and team knowledge building, internalized and externalized knowledge, and team problem-solving outcomes. *Individual knowledge building* occurs when an individual processes data and incorporates it into his or her knowledge base.

This process may involve reading task-relevant information or interacting with task-relevant technology. *Team knowledge building* involves the transformation and dissemination of individual knowledge into actionable team knowledge. *Internalized team knowledge* describes the knowledge each member holds individually, while *externalized team knowledge* describes relationships constructed from knowledge and the task-relevant concepts the team has established (e.g., through verbalizations and/or the creation of task-relevant artifacts). *Team problem-solving outcomes* are influenced by interactions among team members and whether these interactions contribute to fulfillment of critical task requirements (Fiore, Rosen, et al., 2010). Teams with effective CPS strategies engage in parallel and iterative processes where they synthesize these components in service of constructing knowledge, understanding the problem, and evaluating possible solutions (Fiore et al., 2008).

Various empirical efforts have provided evidence for this model of CPS, including examination of communication logs from experienced teams performing tasks in domains such as aerospace defense command, air operations, and unmanned aerial vehicle planning (Hutchins & Kendall, 2010); retrospective accounts of a complex problem faced by experts in NASA's Mission Control Center (Fiore et al., 2014); and by examining transcriptions of CPS during laboratory studies (Rosen, 2010; Seeber, Maier, & Weber, 2013; Wiltshire, 2015). Two aspects are consistent across all these studies: a focus on the study of team communications, whether from actual dialogue or retrospective accounts, and a prevalence of team knowledge-building processes. Thus, CPS is primarily studied through examination of verbal communications (cf. Keyton, Beck, & Asbury, 2010), and the present study is consistent with this. However, rather than focusing on the identification of critical team knowledge-building processes or their relation to performance, we provide a method for examining if and when transitions between problem-solving phases have occurred.

1.2. Problem-solving phases

Problem-solving phases are defined as "qualitatively different subperiods within a total continuous period of interaction in which a group proceeds from initiation to completion of ... problem solving" (p. 485, Bales & Strodtbeck, 1951). In the macrocognition in teams model there are four problem-solving phases: knowledge construction, team problem model, team consensus, and evaluation/revision (Fiore, Rosen, et al., 2010). The *knowledge construction* phase involves the identification of relevant domain information as well as the development of individual team member understanding of the task and the problem. The *team problem model* phase involves the coalescing of knowledge in which the team begins to develop a shared understanding of the problem. The *team consensus* phase is when the team works to achieve agreement among viable solutions to the problem. Lastly, the *evaluation/revision* phase involves analyzing, testing, and validating the agreed-upon solution against the goals/requirements of the task.

These phases have a strong theoretical basis derived from empirical work on problem-solving (e.g., Bales & Strodtbeck, 1951). But initial work to identify these phases has primarily relied on research methods that involve human raters to identify phases (Biron, Burkman, & Warner, 2008). During CPS, particular communicative processes can overlap across phases, but the relative occurrence of each process should vary as a function of which CPS phase the team is in (Fiore, Rosen, et al., 2010). For example, during the *knowledge construction* phase, we expect to see a prevalence of information and knowledge exchange communications. During the *team problem model* phase, we would expect to see a prevalence of knowledge exchange communications and simple agreement, but in the *team consensus* phase, we would expect to see a prevalence of option generation and goal/task orientation. In the *evaluation/revision* phase, we would expect a prevalence of solution evaluation, situation updates, and reflection communications (see Table 1 for more on these communication processes).

Transitions between problem-solving phases may appear continuous or discontinuous. While the phases may occur sequentially, the reality is that CPS is complex, dynamic (Fiore, Smith-Jentsch, et al., 2010), and includes processes that occur at different timescales (e.g., Amici & Bietti, 2015; Wiltshire, 2015). It is therefore likely that during CPS, teams may cycle through these phases in sequence, switch back and forth between phases, or even follow no apparent sequence. This makes identification of phases based on the relative frequency of communication processes difficult (cf. Bales & Strodtbeck, 1951). Thus, we adopt a dynamical systems approach in which phase transitions have been a critical area of inquiry for understanding the coordinated behavior of systems.

1.3. Dynamical systems and phase transitions

Studies of human communication and dialogue, in a broad sense, have increasingly recognized the importance of dynamical systems approaches for understanding changes in patterns of interaction (Angus, Smith, & Wiles, 2012; Dale, Fusaroli, Duran, & Richardson, 2013; Fusaroli, Rączaszek-Leonardi, Tylén, 2014). This type of approach has been adopted by some researchers studying team performance (Gorman, Amazeen, et al., 2010; Gorman, Cooke, et al., 2010; Likens et al., 2014; Strang, Funke, Russell, Dukes, & Middendorf, 2014) as well as in close relationships and family interactions (Butler, 2011). A dynamical system, in the simplest form, is one that changes over time. But there are often many components in dynamic systems that interact and contribute to the patterns of change that characterize a system's behavior. Stable patterns of interaction emerge as a function of the constraints that are placed upon various components by the other components in the system. In this way, dynamic systems are thought to self-organize such that they maintain coordinated performance (e.g., Richardson, Dale, & Marsh, 2014).

Some prior approaches to complex CPS use terminology that is synonymous with a dynamical systems approach and call for corresponding analytic techniques to be utilized in the study of CPS (see e.g., Fiore, Smith-Jentsch, et al., 2010). Leveraging dynamical systems theory *and* its analytical techniques, we argue, provides the necessary foundation for identifying problem-solving phase transitions. In dynamical systems theory, a *phase transition* is a qualitative shift from one coordinated pattern of behavior to another that is

Table 1
Codes for team communication collaborative problem-solving processes (adapted from Rosen, 2010)

Process	Code	Brief Description
Team information exchange	1. Information provision (IP)	Utterances containing facts about the task environment or situation—simple information that can be accessed from one source in the display and "one bit" statements
	2. Information request (IR)	Question utterances asking for a response of simple information about the task environment or situation, or questions asking for repetition of immediately preceding information
Team knowledge sharing	3. Knowledge provision (KP)	Statements about the task environment or situation that provide either (a) an integration of more than one piece of simple information, or (b) an evaluation or interpretation of the meaning, value, or significance of information with regard to the current subtask
	4. Knowledge request (KR)	Question utterances that request a complex information response about the task environment or situation: to answer the question, the response should provide either (a) an integration of more than one piece of simple information, or (b) an evaluation or interpretation of the meaning, value, or significance of information within the current subtask
Team solution option generation	5. Option generation—Part (OG-P)	Statements that provide an incomplete solution—a sequence of actions (i.e., getting a certain tool) intended to contribute to a given subtask—or ask for further refinement and clarification of a solution. These are propositional and suggestive in nature
	6. Option generation—Full (OG-F)	Statements explicitly proposing a complete or near-complete solution—a sequence of actions intended to accomplish part of the task. A complete solution includes reference to specific actions, tools, system components, and actors
Team evaluation and negotiation of alternatives	7. Solution evaluation (SEval)	Utterances that (a) compare different potential solutions, (b) provide support, criticism, or indifference to a potential solution, or (c) ask for evaluation of a solution

(continued)

Table 1 (continued)

Process	Code	Brief Description
Team process and plan regulation	8. Goal/task orientation (GTO)	Utterances directing the team's process or helping it do its work by proposing questioning, or commenting on goals for the team or specific actions team members need to take to address a goal. These statements direct what the team should do next or later in the future. This includes self-references for an individual and are generally more assertive and focused on individual tasks
	9. Situation update (SU)	Statements that provide information regarding what the team is currently doing or what is currently happening with the simulation
	10. Situation request (SR)	Statements that ask about what the team is currently doing or what is currently happening with the simulation
	11. Reflection (R)	Utterances that provide or ask for a critique or evaluation of the performance of the team as a whole or of individual members
Other	12. Simple agree/disagree/ acknowledge (S)	Simple agreement/disagreement utterances are expressions of agreement or disagreement with no rationale provided. Acknowledgements are utterances providing recognition of receipt of communication
	13. Incomplete/filler/ exclamation (INC/F/EX)	Fillers are sounds or words that are spoken to fill gaps between utterances. An exclamation is an utterance that has no grammatical connection to surrounding utterances and emphatically expresses emotion such as laughter. Incomplete utterances are statements that have no explicit meaning because they are missing one or more critical components of grammar: subjects, verbs, or objects
	14. Tangent/off-task (T/OT)	Non–task-related statements, including jokes, sarcastic comments, comments on the nature of the experiment, and statements that have nothing to do with the task at hand
	15. Uncertainty (UNC)	Uncertainty statements explicitly express either general or specific uncertainty about the roles, tasks, situations, or anything else task-related

often, but not necessarily, abrupt (Kelso, 2009). Here we can see the similarities between the definitions of phase transitions. Importantly, as a dynamical system approaches a phase transition, the constraints among the system's components begin to break down. A number of methods have been proposed as a means of detecting these phase transitions, such as changes in variability and autocorrelation (Dakos, Van Nes, D'Odorico, & Scheffer, 2012), examining shifts in the values to which the system is attracted (Butler, 2011), changes in entropy and power law scaling (Stephen, Boncoddo, Magnuson, & Dixon, 2009; Stephen & Dixon, 2009; Stephen, Dixon, & Isenhower, 2009), measures of fluctuation (Schiepek & Strunk, 2010), and, more qualitatively, cognitive event analysis (Steffensen, Vallée-Tourangeau, & Vallée-Tourangeau, 2016).

In the current study, we focus on *entropy*, which is an information theoretic quantification of the amount of disorder/complexity in the system as a function of the number of bits of information needed to describe that system (Shannon & Weaver, 1959). Entropy can serve as an indicator that a dynamical system is undergoing a phase transition because it is inversely related to the number of constraints among the components of a dynamical system (Kugler & Turvey, 1987). Further, we focus on entropy specifically because "it is possible to observe important transitions in systems' behavior by tracking the level of entropy across time or experimental conditions" (p. 215; Guastello, 2011a). Phase transitions can occur abruptly due to a rapid injection of external forces to the system (e.g., change in task demands) or more continuously due to the slower time-scaled natural interaction of the system's components (e.g., Hinrichsen, 2000). As a dynamic system approaches a phase transition, and the constraints of the system begin to break down, the system will exhibit more disorder in its behavior (i.e., higher entropy). Thus, if one examines a continuous entropy time series, peaks in entropy are claimed to be indicative of the system undergoing phase transitions (Stephen, Boncoddo, et al., 2009; Stephen, Dixon, et al., 2009).

Although phase transitions have been argued to be pervasive in social systems (Levy, 2005) and ubiquitous in physical systems (Hinrichsen, 2000), few have studied them in a problem-solving context. In one of the few studies to examine this, the focus was on an individual's ability to solve gear rotation problems (Stephen, Boncoddo, et al., 2009; Stephen, Dixon, et al., 2009). The authors investigated the dynamics of individuals' eye movements and found that, initially, participants would visually trace the rotation of the series of gears in order to determine which way the final gear rotated. But, after a series of trials, participants discovered a new way to solve the problem by counting the number of gears and determining the rotation direction as a function of whether the number of gears was odd or even. This discovery constitutes a problem-solving phase transition that corresponds with a rise and fall (i.e., a peak) pattern in entropy at the transition point (Stephen, Boncoddo, et al., 2009; Stephen, Dixon, et al., 2009).

While not examining phase transitions explicitly, researchers from the interactive team cognition approach have examined entropy in the context of team performance. Specifically, Stevens and colleagues (e.g., Stevens, 2012; Stevens et al., 2012; Likens et al., 2014) have attempted to understand how entropy, estimated from the distribution of team neurodynamic states, relates to their expertise and performance. Entropy in this context corresponds to the degree of cognitive flexibility exhibited by the team, with higher entropy corresponding to more cognitive flexibility. That is, teams with higher entropy are utilizing more of the potential states available to them and thus are more flexible than teams utilizing only fewer states. Specifically, Stevens (2012) found that teams who performed better on team tasks exhibited higher entropy than teams that performed more poorly. Importantly, though, entropy of team states should not be indiscriminately high because it would then reflect a system organization that is random. Thus, while higher entropy appears to be associated with better team performance, for it to have such a relationship it must also be significantly less than entropy for random sequences of states. While this research group has emphasized the importance of entropy for capturing meaningful transitions in team task contexts, they have not explicitly examined entropy in terms of phase transition points or examined collaborative problem-solving, specifically.

1.4. Overview of the present study

For our study of CPS, we examine team communications during a CPS task using a team knowledge-building coding scheme derived from the macrocognition in teams model (Rosen, 2010; see Table 1). These codes represent semantic categories of CPS processes (i.e., information provision, solution option generation, etc.). The sequential concatenation of these codes, derived from observation of a CPS task, form a discrete, nominal time series (Gorman, Cooke, et al., 2012; Gorman, Hessler, Amazeen, Cooke, & Shope, 2012). As such, to derive a continuous measure of entropy on this type of time series, a windowed entropy procedure was used (Gorman et al., 2016; Likens, Amazeen, Stevens, Galloway, & Gorman, 2014). This type of procedure quantifies the entropy within a given window size to characterize disorder of the distribution of states in that segment of the time series. By sliding the window, one is able to make relative comparisons on the basis of how the entropy characterizes the distribution of states from one window to the next. For example, when a team exhibits relatively lower entropy, the distribution of their communication states is relatively stable/orderly. Conversely, when a team exhibits relatively higher entropy, the distribution of states is changing/less stable (cf. Gorman et al., 2016).

This work represents an essential first step in providing evidence that phase transitions occur during CPS and a method for identifying when those points occur. We also sought to determine if there was a relationship between entropy at the peak points and CPS performance. In the current study, we use team communication data taken from a larger study examining team interaction dynamics during CPS (Wiltshire, 2015), and apply a sliding window entropy technique to derive a continuous measure of entropy for each team over the duration of the CPS task. We then specify several criteria for identifying peak data points and assess the robustness of peaks by applying a moving average smoothing technique. From this, we are able to determine the location and relative proportion of entropy peaks within teams.

We predicted that there would be a consistent entropy peak pattern across teams' communications time series. Further, because we expected phases to be distinct, we also predicted that the relative occurrence of codes would vary significantly by phase. The combined evidence of a robust peak pattern and distinct distributions of communication codes due to phases would provide compelling evidence that teams exhibit phase transitions during CPS. Additionally, we predicted there would be a positive correspondence between entropy at peak points and CPS performance, with peak point entropy also being significantly lower than entropy peaks from randomized communication sequences.

2. Method

This work was part of a larger study (Wiltshire, 2015) for which we selected a complex CPS task: NASA's Moonbase Alpha simulation (NASA, 2011). Our study utilized a single interaction design (cf. Malloy & Albright, 2001) because it allows for the natural emergence of interaction dynamics within the constraints of the CPS task (Louwerse, Dale, Bard, & Jeuniaux, 2012).

2.1. Participants

A total of 86 undergraduates from a large southeastern United States university voluntarily participated in this experiment. They comprised 43 dyadic teams. To be included in the present study, participants must have had general video game experience using a mouse and keyboard for third-person video games, no prior history of seizures, no experience using the Moonbase Alpha simulation, and no prior acquaintance. However, to be included in the present analyses a minimum communications time series length of 100 time points was necessary (i.e., there were at least 100 communication-coded utterances). Due to violation of this criterion, a total of three teams' data were excluded from further examination equaling a total of 40 dyadic teams (80 total participants) included in the present research (31 female, M_{age} = 19.2 years, range 18–28 years; ~67% White, 8% Black, 10% Hispanic, 10% Asian, and 5% Other). There were four female-only teams, 17 male-only teams, and 19 mixed-gender teams.

2.2. Materials

Two desktop computers were set up so that participants sat face to face with each other and the computers were offset slightly to one side. This setup also allows for the computer screens to be placed back to back such that participants cannot view each other's screens, which strengthens the need for speech to facilitate collaboration. Each participant was required to wear a Cyber Acoustics AC-840 Internet Communication USB monaural headset with boom microphone to record a single-channel audio file of each participant's communication during the experiment. Audacity was use to record the audio files.

2.2.1. Moonbase Alpha task

NASA's Moonbase Alpha is a complex, CPS task (NASA, 2011). This simulation places team members in a scenario where a meteor strike damages critical life support systems of a moonbase, and they must collaboratively solve the problem of repairing and restoring critical components of the system using a variety of tools to restore oxygen. The goal of the Moonbase Alpha task is for participants to fix and/or replace damaged components of the life support systems to fully restore oxygen to the settlement in 25 min or less. The major components that require repair include solar panels, power cables, couplers, a power distributor, and the life support system itself. A variety of hand tools, robots, and coordination strategies must be employed to complete the task. There are no predefined sequences or guidelines for how to completely repair the settlement in the given timeframe; however, some strategies are better than others.

This problem-solving task requires the collaborative efforts of both team members. That is, given the complexity of the task and the required subtasks, there is a very low probability of completing the mission in the 25 min time limit without collaborating with each other. Further, collaboration is also necessary because of constraints placed on participants by the task. For example, an astronaut can only handle one object at a time. To repair the power coupler and restore energy flow requires not only welding the power coupler itself, but also connecting the power cables, and then tightening the cables with a wrench. The amount of time it can take to handle these three objects can quickly add up given the aforementioned constraint. By contrast, if participants collaboratively delegate that each participant will carry a different tool to complete a different aspect of this task, then a significant amount of time would be saved. Also, because participants do not have any practice with the controls for the task, other nuances and particularities for controlling the astronauts and completing repairs can be communicated to improve efficiency. Thus, communication is essential to successfully completing the task within the time constraints.

2.3. Team communications time series

Team communications were transcribed and coded to represent team knowledge-building processes using a coding scheme adapted from Rosen (2010). Descriptions of the codes are shown in Table 1. The coding process was highly reliable (upon completion of coding $\kappa = .71$; see Appendix A for all coding details). Coded communications form the basis for the discrete sequential time series representing the team's overall communications. This means that each code represents a discrete category (see Table 1) and is listed sequentially as it occurred during the interaction (e.g., Gorman, Cooke, et al., 2012). To create a time series from categorical codes, each code is assigned a value as shown in Table 1. The coding scheme had a range of values spanning 1–15. Through this coding process, a total of 40 teams' time series were included in the present analyses. The range of time series lengths was 102–638 coded utterances ($M = 294.87$, $SD = 137.05$).

2.4. Performance

Problem-solving performance using the NASA Moonbase Alpha simulation was determined as a rescaled combination of three variables: (a) the total time taken to restore life support (0–25 min), (b) the total percentage of oxygen restored (0–100%), and (c) a ratio of completed object repairs to the total possible repairs (only for teams that restored zero oxygen). Because the task was complex, and required both team members to collaborate in order to complete it, some teams did not even complete the mission in time. Therefore, this performance measure accounts for those who not only completed the task with varying durations of time, but also those who did not complete the task yet still performed to varying degrees toward task completion. In other words, the performance measure allows quantification of teams' problem-solving outcomes, even when they do not complete the task, but vary either in their progress toward oxygen restoration or in the number of objects repaired.

The ratio for object repairs was calculated by determining the minimum number of repairs and replacements for each component of the moonbase required to complete the task (25) and then observing each team member's video recording and determining how many repairs and replacements each dyad made. The total completed object repairs/replacements were then divided by the total number of possible repairs/replacements to comprise the aforementioned ratio.

The rescale function in R (R Core Team, 2014) was used to place teams whose performance restored no oxygen at all into a range of 0–33 as a function of their ratio of object repairs/total possible object repairs. Those teams that restored some, but not all, oxygen were rescaled to fit a range of 34–66. Lastly, for those teams that restored all oxygen, the time to complete the task was inversely rescaled to fit the range of 67–100 (with lower times leading to higher scores).

2.5. Gender composition

We used a series of dummy codes to indicate whether the dyad was a female-only dyad, male-only dyad, or mixed-gender dyad. The mixed-gender dyad was used as the referent group.

2.6. Knowledge assessment

A custom 10-item multiple-choice knowledge assessment was developed as a way to control for individual differences in knowledge acquired from the Moonbase Alpha training presentation given to participants. The values for this measure were computed by summing the total number of items answered correctly with a possible range of 0–10. These items reflected important details of the training, such as how to repair damaged components of the life support system, how to use certain tools, and what types of information should be communicated between team members. We adopted a *selected score model* approach (Chen, Mathieu, & Bliese, 2004), which suggests an appropriate, aggregate level variable representing the team from individual variables for each member of the team. We assessed the strength of relationship between the mean (i.e., the average of the two team members' scores), minimum (i.e., the lower of the two team members' scores), and maximum (i.e., the higher of the two team members' scores) of the knowledge assessment scores for each team. We determined that the minimum knowledge assessment score had the strongest correlation with performance ($r = .41$, $p < .05$); therefore, this served as our selected score to represent dyads' collective task knowledge. The full list of items for the knowledge assessment can be found in Wiltshire (2015).

2.7. Procedure

Upon arrival at the laboratory, participants were briefed about the nature of the experiment and then asked to introduce themselves to each other by providing a greeting and sharing their name with the other participant. Participants were then given an informed consent document to review and asked to complete a biographical questionnaire.

Participants were then given a PowerPoint tutorial that covers the basics of the Moonbase Alpha simulation and the problem-solving task. The information presented to participants was derived from the simulation's instruction manual (NASA, 2011). Further, participants were told that they would be tested on the content in the PowerPoint. After each participant completed the PowerPoint, they received a short 10-item multiple-choice knowledge assessment.

After completion of the knowledge assessment, the necessity for communication to complete the task was reiterated. Participants were then prompted to put on the audio headsets and instructed to begin the simulation. A short introduction video conveyed the nature of the problem (i.e., the moonbase was damaged by a meteorite and life support functions need to be restored) before participants began the 25-min task. The Moonbase task was considered completed either when time ran out or once participants fully restored oxygen to the moonbase, whichever came first. The study was concluded by asking participants if they had any questions about the nature of the research, answering any questions they had, and providing a research evaluation form.

2.8. Sliding window entropy

In order to derive a continuous entropy time series, we calculated a sliding window entropy on each teams' communications time series. Shannon information entropy (Shannon & Weaver, 1959), one of the most widely examined forms of entropy, can be thought of as an indicator of the amount of order versus disorder exhibited by a system (see Guastello, 2011a, for a detailed discussion). Shannon information entropy assumes there is a set of discrete states a system can exhibit and that there is a probability associated with the occurrence for each state.[1] The equation for Shannon information entropy in this context is:

$$-\sum_{i=1}^{15} p_i \times \log p_i \tag{1}$$

In this equation, p_i corresponds to the relative probability that a given communication code i occurred, where i is an indicator of one of the 15 communication codes characterizing the team knowledge-building process (see Table 1). To better understand the meaning of this calculation, consider the following example. Assume that there is a communication series where there are only two possibilities: knowledge request and knowledge provision. If there were a case where the probability was 1 for a single communication code (e.g., knowledge request), then there would be a probability of 0 for all other codes. If we were to plug these values into the equation we would get a very low or 0 value. There is no variability in the sequence of values and it is, therefore, highly ordered. Recal, that higher entropy is associated with more disorder and lower entropy coincides with more order. Now consider a case where there is an equal probability of observing any one of the communication codes. The entropy value would be maximized, representing very high disorder, and perhaps randomness, because we can say with very little certainty that one state or another will be exhibited.

When we assess a communication sequence with the Shannon information entropy algorithm, we get a single entropy value for that entire sequence. At issue is that the ordering and distribution of the states in the communication sequence changes over the duration of a task (e.g., Gorman, Hessler, et al., 2012). A sliding window calculation of entropy can capture fluctuations in the patterns of states that the system exhibits over time. It provides a series of entropy estimates from each window with a specified size, with a specified iteration interval. In the current case, Shannon entropy was calculated for a window size of 25 and iterated by one time point. This results in a continuous entropy time series that is of length N-24 (time unit offset can be seen by comparing the *x*-axes of the top and bottom graphs in Fig. 1). By using this window size and iterating by one time point, we get the relative contribution of each new time point (cf. Gates, Gatzke-Kopp, Sandsten, & Blandon, 2015; Gorman et al., 2016), and, thus, a continuous measure of entropy that fluctuates over the duration of the task as a function of the ordering of communication sequences. See Appendix B for a schematic illustrating the relationship between the communications transcription, codes, and sliding window entropy.

We determined the appropriate window size by calculating the average mutual information for each time series and identifying the lag of the first local minima (Fraser & Swinney, 1986) using R (R Core Team, 2014). We found the average first local minima across all of the time series was at a lag of 25, although minima ranged from a lag of 2 to 94. This suggests that informational dependence is greatly reduced on average at 25 time points forward. We also examined the entropy plots for a range of window sizes from 5 to 100. Some have noted that small window sizes are subject to artefactual spikes (Gorman et al., 2016; Likens et al., 2014) and biased entropy estimates (Bonachela, Hinrichsen, & Munoz, 2008). However, based on examination of the entropy plots at various window sizes, we found a window size of 25 captured dynamic changes in entropy

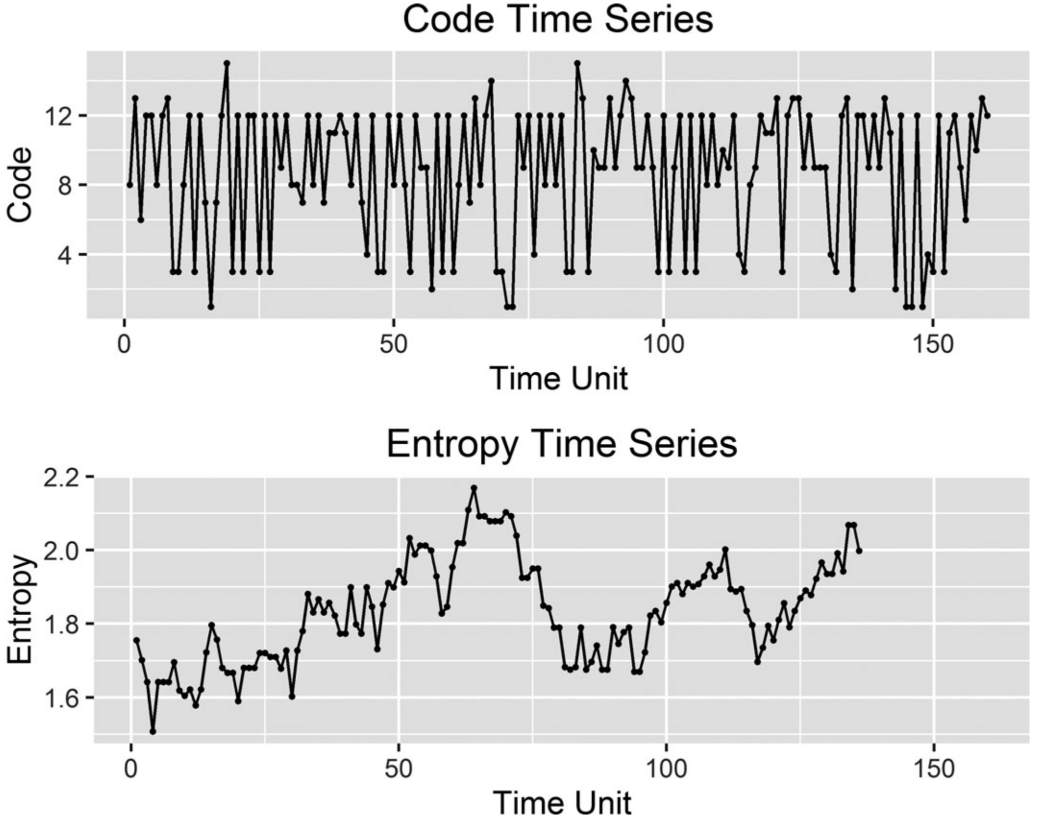


Fig. 1. Example of communications code time series (top) and continuous entropy time series (bottom).

without extensive spiking or smoothing (see Fig. 1 bottom; note, however, that we later apply smoothing to increase the robustness of peak identification). This window size is also more likely to capture intrinsic, as opposed to extrinsic, fluctuations in the communication patterns (Gorman, Hessler, et al., 2012).

2.9. Identifying peaks in entropy time series

We developed an algorithm to detect peaks in the time series by capitalizing on time-delay embedding (see Boker & Laurenceau, 2006). First, we generated a lead and a lag from the observed entropy time series for each team with a time delay of $\tau = 1$. We truncated all time points that generated missing data (i.e., the first and last time point of each time series). We then specified two criteria that must be satisfied in order for a given point to qualify as a peak. The current entropy value must be higher than the lead entropy value ($\tau - 1$), and also higher than the lag value ($\tau + 1$). Using these two criteria, we generated a binary time series that indicated whether a given point satisfied these two criteria, with a value of 1 indicating the presence of a peak and 0 reflecting the failure to meet these criteria.

Thus, using the two criteria, we identified the time points where there were peaks in entropy across all teams' communications time series. We then calculated the *peak proportion* within each team by dividing the total number of peaks within a given team's time series by the total length of that time series. Upon examination of several time series, we observed several featured instances where points that were identified as peaks, based on our criteria, but would not have been visually identified as peaks (see Fig. 3 top). These points are problematic because they often fell at a point that should not be considered a peak (e.g., a valley at time point 50 in Fig. 3 top). Thus, we sought to improve the peak identification technique.

In order to increase the robustness of our peak identification, we applied a moving average smoothing algorithm to the observed entropy time series of each team. We tested a variety of window sizes ranging from 2 to 15 to observe how the average *peak proportion* changes as a function of the moving average window size. We also calculated the 95% confidence intervals of the mean for each of the window sizes. In Fig. 2, we show the mean peak proportion and 95% confidence intervals (y-axis) for each of the window sizes (x-axis). The value of 1 on the x-axis represents the original peak proportion values without any smoothing. In order to identify the appropriate amount of smoothing, we selected the first moving average window size that had a prior and subsequent window with an overlapping confidence interval. A reason for selecting this lower value is that the larger the moving average window size the more data are lost. From Fig. 2, it can be seen that this is at a window size of 5. There is a large drop in the peak proportion from the original entropy time series to a window size of 4 (~6%) and that begins to level off with each increasing time point added to the window. Again, applying the moving average smoothing technique here ensures that the identified peaks are robust and remain

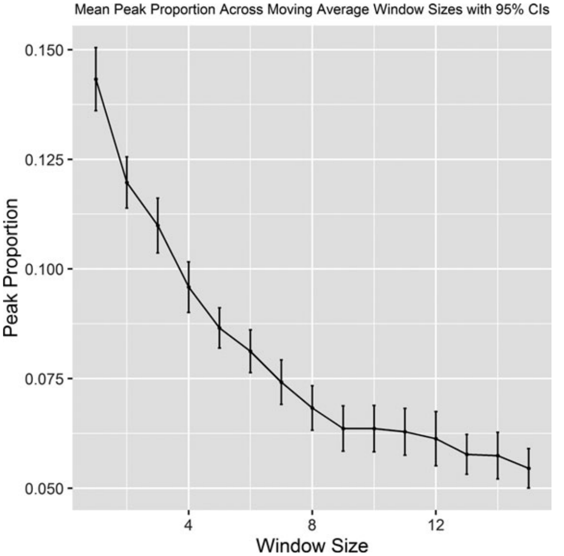


Fig. 2. Mean peak proportions at different moving average window sizes.

peaks following the smoothing. Thus, all non-robust peaks fail to meet the peak criteria due to the smoothing process.

Fig. 3, bottom, provides a visualization of how our criteria for identifying peaks maps to the observed entropy value. Vertical lines represent time points that our criteria identified as peaks. Note the reduction in non-robust peaks when comparing the top and bottom graphs. As support for the validity of our peak identification procedure, it is clearly apparent that the majority of peaks in the bottom graph of Fig. 3 would be visually classified as peaks. This pattern (i.e., comparison of invalid peaks before smoothing to valid peaks after smoothing) was consistent for all of the time series.

By knowing which data points were peaks and their respective time unit in the smoothed entropy time series, we can trace these points back to the original code communications time series. First, recall that the window size (25) and step (1) for the sliding window entropy gives us a continuous entropy time series of length N-24. Next, after applying smoothing to the data with a moving average window size of 5, we will have a length of N-29 of the original time series. Thus, it is relatively straightforward to identify

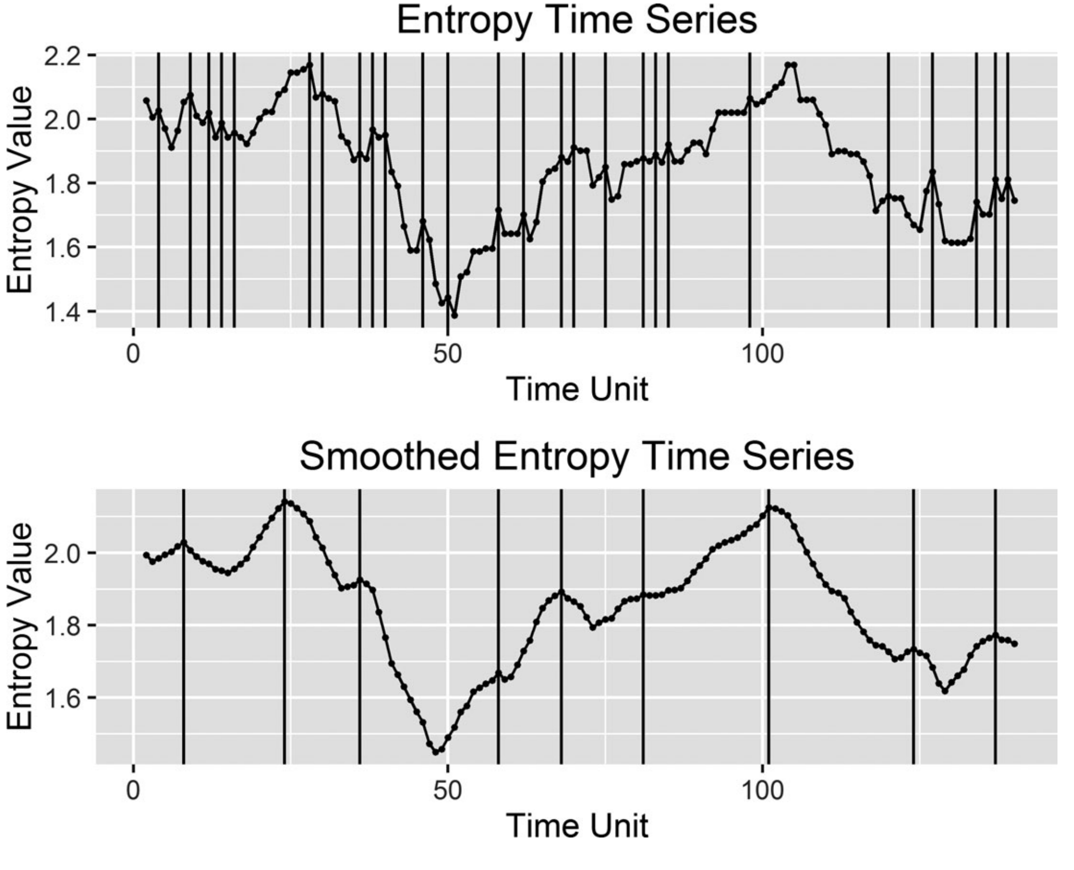


Fig. 3. Original (top) and smoothed (bottom) entropy time series for an exemplar team with vertical black lines representing peaks identified with our criteria.

peaks in the original time series with the caveat that we cannot identify peaks in the original time series that fall within the last 29 time points. This, of course, varies depending on whether a different window size was used for estimating entropy and/or for smoothing with the moving average.

Using this information, we generated a new variable for each of the original time series that we label *Epoch*. To do so, we first truncated the last 29 time points from the time series. Then, we used the binary peak point variable to determine where to define the epochs (epochs began/ended at points we determined to be peaks, for which the variable equaled 1). All values up until the first peak point were assigned a value of 1, and then all values up until the next peak point were assigned a value of 2, and so on. In this way, we generated a clustering variable that indexes, in increasing order, the epochs that segment the time series as a function of being bounded between peaks in entropy. This was done for each team's time series. Note that we refer to these segments as epochs prior to verifying they are actually distinct phases. R scripts and SPSS syntax for the peak identification procedure described above and most of the analyses below are available at https://github.com/travisjwiltshire/PhaseTransitionsDuringTeamCollaboration.

2.10. Analytic strategy

Descriptive statistics of the entropy peak data were generated using R. All other analyses reported below were conducted in SPSS 21 (IBM Corp. Released 2012. IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp. USA). In order to test the hypothesis that the distribution of codes varies by phases, we estimated a series of multilevel models with full maximum likelihood using each communications code as a dependent variable; that is, a separate variable was generated to indicate the frequency of each communications code. We rescaled each of these variables to have a range of 0–100. We ran intercept-only multilevel models that included random effects for the ID variables Team (a numerical team identifier) and Epoch (a numerical identifier increasing in value that specifies the clustering and order of communication segments nested within Team).

Further, due to the fact that the time series lengths for each team and the length of each epoch could vary, we included a weighting of one over the number of units in the epoch so that time series of different lengths contribute equally (Cohen, Cohen, West, & Aiken, 2013).

Our analyses can be thought of as partitioning the variance in the prevalence of each code that can be attributed to differences due to team as well as due to epoch within team. These analyses are a direct analog to the decomposition of variance utilized in ANOVA models which are based on the factorial sums of squares decomposition (Shavelson & Webb, 1991; Tabachnick & Fidell, 2013). From this, we can explain variability in code prevalence. These analyses tell us whether the prevalence of given code varies as a function of the Team and Epoch ID variables. We view this as a significant improvement upon previous approaches that used arbitrary time segmentation for epochs and repeated chi-square tests for each communication code, each epoch, and each group (e.g., Hirokawa, 1983).

Following these analyses, we calculated two new variables: observed average peak point entropy and randomized average peak point entropy. The *observed average peak point entropy* was created by taking the average of the entropy values from the peak points of the smoothed time series only. The *randomized average peak point entropy* was conducted by following the same procedure for identifying peak points and calculating the observed average peak point entropy. However, instead of applying this to the observed communications time series, they were based on 40 randomly generated code sequences (drawing from a pool of values 1–15 with replacement) that had random length between 102 and 638, matching the length range of the sample exactly. The randomized average peak point entropy reflects the amount of entropy at the peak points expected due to chance. This is akin to surrogate data testing that is commonplace in examining entropy-based and other dynamical measures in the literature (e.g., Ramseyer & Tschacher, 2010; Stevens, 2012; Strang et al., 2014) and allows for determination that the observed average peak point entropy reflects a meaningful amount of order/disorder.

Next, we conducted a hierarchical multiple linear regression to examine the relationship between observed average peak point entropy and performance. We controlled for variability in performance due to the gender composition of teams and knowledge assessments scores by entering these variables into the first step of the model. Then we entered observed average peak point entropy in the second step. We also wanted to ensure that observed average peak point entropy was less than randomized average peak point entropy, so we compared them using a paired samples t test. We also tested whether there was any relationship between performance and the randomized average peak point entropy.

3. Results

The range of observed entropy values across all teams was 0.17–2.45. When collapsing across all teams, there was a total of 1,519 data points that met our two peak criteria prior to smoothing. With a total of 10,683 data points, approximately 14.22% could be classified as peaks in entropy. When examining the peak proportion within each team's time series, we observed a range of 10.41%–21.64%.

After smoothing with a moving average window size of five, the average peak proportion across all teams was 8.49% (SD = 1.47%). The actual frequencies of peaks observed within each team ranged from 3 to 51 peaks with an average of $M = 22.12$ peaks ($SD = 12.12$). Thus, all teams exhibited robust peaks in their respective entropy time series.

Table 2 shows the results of our multilevel models examining the random effects of the Team and Epoch ID variables on the prevalence of each code. The table includes the decomposition of variances that are the mean contribution by Team, Epoch, and residual error. Effect sizes were calculated by taking the estimates for a given random effect and dividing it by the total variance including the residual. This can be interpreted like a traditional R^2 using ordinary least squares regression as the amount of variance explained by differences due to a given Team or Epoch ID.

Almost all Team effects were significant for each code. When significant, these values suggest that the prevalence of the respective code significantly varied as a function of Team. In other words, some teams exhibited certain communication processes (i.e., codes) more/less than others. Only Option Generation-Full and Reflection did not vary significantly across teams.

All of the Epoch effects were significant for each code. This suggests that the relative occurrence of each code significantly varied as a function of the epoch in which it resided. In other words, the communications codes were distributed differently as a function of the epoch segmentations introduced based on peaks in entropy values. Importantly, when examining the effect sizes for Epoch, a substantial amount of variability can be accounted for (i.e., nearly half) in the distribution of each code across epochs. Thus, because the occurrence of codes varied significantly across epochs, we can interpret them as qualitatively distinct CPS phases.

Taken together, and consistent with our predictions, these findings suggest that teams do exhibit phase transitions in their CPS communications due to the two findings: (a) We observed that the entropy peak pattern remained robust across all teams and (b) that phases identified using entropy peaks phases were distinctly variable across epochs.

3.1. Relation between average peak entropy values and CPS performance

Results from the hierarchical linear regression are shown below in Table 3. As expected, and consistent with correlation analyses, the minimum knowledge assessment scores were significant predictors of performance. The higher the minimum scores on the knowledge assessment for a team, the better they performed. Relative to teams composed of mixed genders, male-only teams were more likely to perform significantly better and female-only teams were also more likely to perform better, but not significantly. Taken together, these variables accounted for approximately 40% of the variability in performance.

After controlling for variability due to gender composition and knowledge, we observed that lower observed average peak point entropy predicted better problem-solving performance. The addition of this variable specifically characterizing entropy at the transition points accounted for an additional 8% of the variance in performance.

As entropy values increase, they move from meaningful disorder of states toward randomness. Thus, to assess the degree to which the observed average peak point entropy values were not random, we compared them with the randomized average peak point entropy. The observed average peak entropy values ($M = 1.96$, $SD = .09$) were significantly lower than the randomized average peak point entropy values ($M = 2.41$, $SD = .05$, $t(39) = -26.97$, $p < .001$). Further, randomized average peak point entropy values were not predictive of performance ($\beta = .04$, $t(39) = -0.29$, $p = .77$).

3.2. Qualitative examination of communication processes within phases

In this section we provide a qualitative examination of one team's communication processes to illustrate the code distinctions across phases. Our aim here is not to be

Table 2
Results from multiple intercept-only random effects multilevel models characterizing variance in a given communications code that can be attributed to Team and Epoch

	Variance	*SE*	*p*	Effect Size
C1. Information provision				
Residual	51.65	0.74	.000**	.47
Team	11.89	3.68	.001**	.11
Epoch	45.08	4.61	.000**	.42
C2. Information request				
Residual	32.59	0.47	.000**	.46
Team	4.80	1.89	.011*	.07
Epoch	32.72	3.12	.000**	.47
C3. Knowledge provision				
Residual	128.17	1.78	.000**	.49
Team	24.10	7.91	.002**	.09
Epoch	106.80	10.97	.000**	.41
C4. Knowledge request				
Residual	66.22	0.96	.000**	.45
Team	4.37	2.22	.049*	.03
Epoch	75.40	6.71	.000**	.52
C6. Option generation-full				
Residual	12.74	0.19	.000**	.39
Team	0.79	0.50	.123	.02
Epoch	19.53	1.53	.000**	.59
C7. Solution evaluation				
Residual	6.93	0.10	.000**	.24
Team	2.04	0.78	.009**	.07
Epoch	19.63	1.25	.000**	.69
C8. Goal/task orientation				
Residual	84.60	1.22	.000**	.46
Team	17.75	6.06	.003*	.10
Epoch	80.56	7.89	.000**	.44
C9. Situation update				
Residual	74.70	1.08	.000**	.48
Team	7.75	3.21	.016*	.05
Epoch	70.13	6.89	.000**	.47
C10. Situation request				
Residual	16.80	0.24	.000**	.54
Team	1.68	0. 73	.025*	.05
Epoch	12.88	1.42	.000**	.41
C11. Reflection				
Residual	11.49	0.17	.000**	.46
Team	0.57	0.42	.0174	.02
Epoch	12.84	1.16	.000**	.51
C12. Simple agreement/disagreement				
Residual	88.94	1.28	.000**	.49
Team	28.46	8.43	.001**	.17
Epoch	64.89	7.37	.000**	.36

(continued)

Table 2 (continued)

	Variance	*SE*	*p*	Effect Size
C13. Incomplete/filler/exclamation				
Residual	108.69	1.57	.000**	.37
Team	48.97	13.37	.000**	.17
Epoch	133.74	11.51	.000**	.46
C14. Tangent/off-task				
Residual	14.29	0.21	.000**	.45
Team	2.57	0.85	.002**	.08
Epoch	14.92	1.39	.000**	.47
C15. Uncertainty				
Residual	19.68	0.28	.000**	.50
Team	1.37	0.69	.046*	.03
Epoch	18.44	1.81	.000**	.47

Note. *p < .05; **p < .01.

Table 3
Results from multiple regression model used to predict collaborative problem-solving performance

Predictor	Beta	*t*	*p*	Adj R^2
Step 1				
Male only	.60	4.63	.000**	.39
Female only	.26	2.00	.053	
Minimum knowledge assessment	.30	2.44	.02*	
Step 2				
Male only	.56	4.59	.000**	.47
Female only	.26	2.15	.039*	
Minimum knowledge assessment	.29	2.51	.017*	
Observed average peak entropy	−.30	−2.51	.017*	

Note. *p < .05; **p < .01.

exhaustive in showing how observed communication phases fit with theory, but rather to provide a comparative example. Future qualitative analyses could concomitantly examine the distributions and transcriptions. For the example team, we selected based on a lower frequency of phases (11 phases). The original communications time series was 179 units long and the smoothed entropy series was of length 150. We observed peak points at times 7, 23, 35, 57, 80, 100, 123, 136, and 145. We used these time points to segment the original time series giving us segments with time units of 2–7, 8–23, 24–35, 36–57, 58–80, 81–100, 101–123, 124–136, 137–145, and 146–150. Thus, this example team had a total of 11 phases. Fig. 4 shows the frequency distribution of each categorical code across the 11 phases exhibited by the team.

Drawing on macrocognition in teams theory, we interpret the distributions of codes across phases in light of their “varied modal frequencies” and their differential

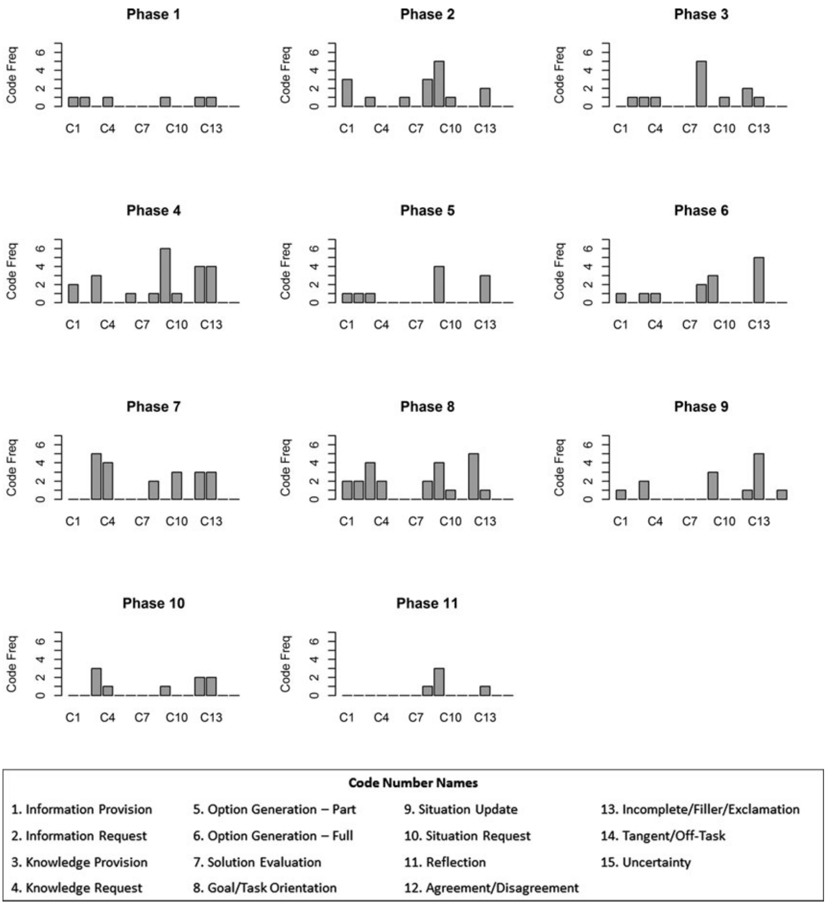


Fig. 4. Example of team's collaborative processes across problem-solving phases.

"occurrence of particular ... processes" (Fiore, Rosen, et al., 2010, p. 215). Phase 1 is characterized by a relatively even distribution of information request and provision, knowledge request, situation updates, simple agreement, and incomplete utterances. In line with the theory, Phase 1 can be considered as an initial phase of *knowledge construction* as the team is working to understand the task. Phase 2 is characterized primarily by information provision (19%), goal/task orientation (19%), and situation updates (31%), as the problem space is built (cf. Fiore & Schooler, 2004). Phase 3 is primarily goal/task-oriented utterances (42%). Phase 4 is primarily situation updates (27%) with simple agreement/disagreement (18%). Phase 5 is primarily situation updates (40%) with some information and knowledge exchange (30%). Phases 2 and 4 are the only phases with instances of option generation and, thus, likely represent team *problem model phases* where the teams are refining their understanding of the problem and working toward effective solutions. It may be that in Phases 3 and 5 the team is exhibiting the *team consensus phase* in which they have agreed on a solution option and are working toward executing that option.

Phase 6 is primarily characterized by incomplete or filler statements, with some situation updates, and knowledge exchange. Phase 7 is primarily characterized by knowledge provisions and requests (45%). Phase 8 exhibits a fairly even distribution of information and knowledge exchange codes, goal/task and situation-related codes, and some simple agreement/disagreement. Phases 7 and 8 may represent a return to the *team problem model* and *consensus* phases, respectively, in which the team is providing new knowledge they have acquired about the problem, setting new goals, and updating their progress toward achieving them. In Phase 9, we observed a high frequency of incomplete statements. In Phase 10, we observe a majority of knowledge exchange codes (44%). And Phase 11 is characterized primarily by situation updates (60%). As this team worked to complete the task, they may have felt time pressure, which may have led them to the issues with incomplete statements in Phase 9, and return to the *team problem model* in phase 10, and updates of the situation in the final phase.

4. Discussion

In the current study, we used a sliding window entropy technique combined with several criteria for identifying peaks to demonstrate that teams exhibited phase transitions in their CPS communication processes. The basis for this argument stems from dynamical systems theory and coordination dynamics. In these areas, entropy has been argued to serve as an indicator that a system is undergoing a phase transition because it is inversely related to the number of constraints among the components of a dynamical system (Kugler & Turvey, 1987). That is, as a system approaches a phase transition, the constraints within the system begin to break down, and the system exhibits higher entropy. Evidence for this peak pattern can be found in continuous entropy time series in which peaks in entropy values are indicative of the system undergoing a phase transition (Stephen, Boncoddo, et al., 2009; Stephen, Dixon, et al., 2009).

This study is unique in that it provides evidence of phase transitions during collaboration without relying on relative frequencies (cf. Bales & Strodtbeck, 1951) or arbitrary time segmentation (Hirokawa, 1983) and focuses on the knowledge-building work characterized by the macrocognition in teams model (Fiore, Rosen, et al., 2010; Fiore, Smith-Jentsch, et al., 2010). In a more general sense, this work also provides additional empirical support for entropy as an indicator of phase transitions, especially for cases where there are multiple peaks in entropy as opposed to a single peak (cf. Guastello, 2011a; Stephen, Boncoddo, et al., 2009; Stephen, Dixon, et al., 2009) and that links entropy at transition points to CPS performance.

In terms of understanding how groups and teams perform effectively in collaborative contexts, identifying points of phase transitions is a critical area of inquiry. In particular, Kelso (1990) noted that "around phase transitions ..., phenomenological descriptions turn to prediction; the essential processes governing a behavioral pattern's stability, change and even its selection can be uncovered" (p. 249). Thus, we view the identification of phase transitions as essential for understanding the coordinative mechanisms facilitating CPS, which can form a solid foundation to further our understanding of team cognition.

Indeed, we found that lower average peak entropy values corresponded to better problem-solving performance. However, this was contrary to our prediction. Teams with lower entropy values at the transition points, and thus more order in their communication sequences at transition points, were more likely to effectively problem solve. There are several substantive explanations to consider here. We based our original prediction on Stevens's (2012) findings that teams who performed better, and had more experience, exhibited higher levels of entropy across the duration of their task.

On the one hand, the two studies share important similarities; while Stevens examined team neurodynamic states rather than team communication states, the general principles should apply in both cases. Whereas higher values of entropy correspond to more of the potential team communication states being utilized within a given window, lower entropy values correspond to fewer of the potential states being utilized. Like Stevens's work, our use of entropy may reflect a spectrum indexing team cognitive stability and flexibility that can be associated with problem-solving performance. Only, in this case, it is an index that is specific to transition points in team communication sequences. On the other hand, the present study diverges from the work by Stevens et al. (2012) in that they used a task that was more focused on decision-making than it was on problem-solving. Further, the participants in that study were not only familiar with the task, but some also had considerable expertise. As such, it could be that experience with the task context alters the nature of the communications, thus modifying the entropy expressed. For example, given the high task familiarity, much of the necessary task knowledge and team coordination may be implicit (Rico, Sánchez-Manzanares, Gil, & Gibson, 2008). This would decrease the frequency of utterances associated with task elements. As such, when phase transitions occur, expert teams may be more likely to begin discussing a more variable set of options in preparation for a new phase. By contrast, with our less expert teams, stability of utterance type within phases may have been the more useful strategy, indicating a "leveling out" of shared understanding that could have produced effective coordination. This would explain why our higher performing dyads exhibited lower transition point entropy. In short, given that entropy-based analyses are a developing area in the team literature, it is important to identify and examine what similarities and differences emerge in the findings and examine why they may be occurring.

It is likely that CPS performance is better when phase transitions capitalize on the optimal amount of communication states. Lower entropy at transition points could reflect a stable shift in team communications that is quite coordinated. By contrast, higher entropy peaks may reflect more drastic transitions. A more general substantive interpretation is that our results may reflect a pervasive phenomenon described by the *theory of optimum variability*. This theory posits that complex systems exhibit effective performance when there is an optimal degree of order and disorder characterizing the system's behavior (Harrison & Stergiou, 2015; Stergiou & Decker, 2011). Thus, effective team performance can be achieved when the right amounts of both stability and adaptability are exhibited (Parker, Best, Funke, Strang, & Marion, 2016), optimally balancing task constraints and resources (cf. Bak, Tang, & Wiesenfeld, 1988; Likens et al., 2014). Given discrepancies between our predictions and results, further research should consider how optimum variability theory informs our understanding of entropy at phase transition points and how it relates to effective team performance.

As an indication of how such findings can be used, during periods of transition, dynamical systems are more easily perturbed by external factors (Butler, 2011). Based upon this, some have noted that periods of transition between phases may be a particularly effective time to apply interventions (e.g., Hollenstein, 2007). Experiments could then vary timing of interventions, based upon phase transitions, to explore differential effects on team process and performance. For example, technologies that prompt and/or elicit the externalization of team cognitive processes when embedded around phases may improve team performance and collaborative cognition (Fiore & Wiltshire, 2016; Gorman, Cooke, et al., 2010; Wiltshire & Fiore, 2014; Wiltshire et al., 2014).

We also found that gender composition and task knowledge were significant predictors of performance. Same-gender dyads tended to perform better than mixed-gender dyads. Also, those teams with higher minimum task knowledge performed better. There are other potential sources of heterogeneity that might contribute to the collaboration, phase transitions, and performance. For example, there could be interpersonal differences in coordination (Abney, Paxton, Dale, & Kello, 2015), dominance (Sadler, Ethier, Gunn, Duong, & Woody, 2009), and communication flows (Fischer, McDonnell, & Orasanu, 2007) that could be investigated in future work.

There are several limitations to note. On the one hand, using entropy alone cannot indicate "what phase" of CPS a team may currently be exhibiting. Rather, it only serves as a metric to indicate that the structure of the system is changing. In other words, a continuous measure of entropy over the duration of a CPS task only captures the fluctuations in the ordering of the system. When a team's communication processes reside within a particular phase, we see that entropy remains low and the system is relatively stable. But, as the distribution of specific CPS processes begins to transition, more disorder is exhibited by the system, which is in turn evidenced by increased entropy leading to a peak. More methodologically oriented future work ought to employ a variety of techniques (see Section 1) for identifying phase transitions in order to assess the convergent validity of entropy peaks as a sufficient marker of transition points.

Concomitantly, there is no satisfying established technique in the literature for how to identify "what phase" is being exhibited. The most widely used technique relies on a judgment based on the most frequent state within a given phase. This presents researchers with a circularly unverifiable signal-to-noise problem of whether a phase transition is accurately being detected by entropy and what the qualitative distinct phases are that are being identified by the changes in entropy. We later discuss potential ways to overcome this problem. Historically, CPS theory has posited four problem-solving phases (Bales & Strodtbeck, 1951; Fiore, Rosen, et al., 2010; Hirokawa, 1983). In our results, we observed an average of approximately 22 robust peaks across teams. This high frequency of implied phases, we think, reflects the nature of this complex CPS task and others like it. For example, while teams are working toward an overall goal of restoring oxygen to the moonbase, numerous subproblems arise during the task such as how to repair specific components of the system that can only be accessed by a robot. Indeed, this *polytelic situation* is a defining characteristic of complex problems (Funke, 2010), and other types of problems, such as insight problems, may exhibit far fewer phase transitions (e.g., Stephen, Boncoddo, et al., 2009; Stephen, Dixon, et al., 2009; Steffensen et al., 2016; cf. Weisberg, 2015). Thus, independently of determining what the phases actually are, the high number of phase transitions implies that this type of complex problem does not follow a clear four-phase solution and, instead, that there is likely cycling and shifting through the various phases (Fiore, Rosen, et al., 2010). Our qualitative example also supports this interpretation.

An additional limitation is that our peak identification criteria do not differentiate between global and local maxima for identifying peaks. Consider Fig. 3 (bottom) where the first peak is much smaller than the subsequent peak or perhaps a large peak has a smaller peak as a part of it. With our current criteria, as long as a given value is larger than its adjacent values, and it remains this way following smoothing, it will be classified as a peak. Future work could identify more relative peak procedures that account for the highest and lowest peaks and could explore the possibility of increasing the number of preceding and succeeding points required to constitute a peak. This may even mitigate the need for smoothing procedures. Additionally, in some of the original entropy time series, we visually observed peaks that were characterized by two points of equal value. The current criteria did not account for this. Although smoothing seemed to mitigate this issue, researchers adopting our method should examine their data for this pattern. Perhaps an additional criterion could be specified that allows for two points of equal value, as long as their adjacent values are lower, to be considered a peak.

We also used the time point of peaks in entropy and demonstrated how these can be markers for time points in the original communications time series. In principle, with the parameters we selected, this could apply to any time point in the time series except for those falling within the last 29 time points. Recall that by using a window size of 25 and a step of 1, we get the relative contribution of each iterative time point (cf. Gates et al., 2015). Therefore, the time points identified as peaks in the entropy time series can be used to identify the phase transition time points in the original time series.

Upon identification of these time points in the communications time series, we found that the relative occurrence of each code varied significantly across phases. We qualitatively examined the distribution of communication processes (i.e., categorical codes) for an exemplar team and noted how they were different prior to and after peaks in entropy time points. In terms of macrocognition in teams theory, the phases identified for this particular team do partially coincide with predicted phases. What was missing was a discussion of solution evaluation/negotiation, which could be due to the nature of the task (e.g., the degree a solution needs evaluation and/or negotiation is related to its complexity as well as ambiguity or disagreement; Fiore, Rosen, et al., 2010). They primarily alternated between information and knowledge exchange, goal/task orientation, and status updates. What is important about this example is that we demonstrated how to visualize the processes within phases in a way that complements our quantitative analyses. Our results show that teams did exhibit distinct phases in their CPS processes, and they did so in a way that they can be qualitatively assessed.

We also suggest additional theoretical refinement is needed in thinking about CPS phases as a dynamical system. For example, the theory suggests a reliance on the modal communication process (i.e., most prevalent code) of a phase to attribute an overall phase description to it (Fiore, Rosen, et al., 2010; Hirokawa, 1983). However, even in the example provided, we can see that these distributions may be multimodal (i.e., multiple codes have high frequencies within a phase). This suggests that theorizing CPS as a dynamical system requires measures more complex than those like central tendency, because they can oversimplify the dynamics of the system. Thus, to enrich macrocognition in teams theory, we suggest a shift in thinking in terms of modes and frequencies to thinking about sequences of processes and their stability during team performance (Araújo et al., 2015).

One way to do this is to adopt techniques that identify the most prevalent sequences of communication processes and their stability within and across sub-distributions (i.e., phases) of the original communications time series. Orbital decomposition analysis is one technique that can identify the most prevalent sequences and their stability, and which has previously been applied in a group problem-solving context (e.g., Guastello, 2000). This analytic technique combines a variety of metrics including stability, information, and complexity to better understand dynamic patterns specifically in categorical time series data (Guastello, 2011b; Pincus, Ortega, & Metten, 2011). Combining the methods used in this paper with orbital decomposition analysis (cf. Pincus & Guastello, 2005) would allow team performance and CPS researchers to better determine whether empirical observations of phases identified using a sliding window entropy coincide with theoretical predictions of CPS phases. Such an analytic framework may lead to theoretical enrichment that allows us to characterize problem-solving phases in terms of prevalent sequences of communication states that are stable within particular phases (and not within others). Such an approach would actually assume that phase identification should not be about a priori frequencies or modes of particular states, but that the emergent patterning of states (i.e., sequences), whatever they may be, should be the defining feature of the phases.

However, if researchers are disposed to rely on distributions of states in phases, then a more systematic way of assigning semantic descriptions to the distribution of codes in each phase is needed. A computational means of developing a pattern classification system may be necessary, similar to that used by Stevens et al. (2012) to categorize different distributions of neurophysiological measurements of engagement across multiple team members. This would provide a more objective classification of code distributions as representing particular theoretical phases that are based more on the dynamics of the CPS processes as opposed to the modal description of those processes.

Upon identification of distinct phases, together with semantic descriptions of what those phases are (based on empirical analysis and observation), the next step would be to determine the relation between sequences of phases and CPS performance. That is, one could determine whether particular sequences of CPS phases relate to how well a team solves a problem. Given the high number of observed phases within time series, we expect that certain iterative cycles of the phases may be optimal for effective CPS performance. However, we make this assertion with some reservations because Hirokawa (1983) concluded that "'successful' and 'unsuccessful' problem solving groups cannot be easily distinguished solely on the basis of the sequences of interaction phases that characterize their problem-solving discussion" (p. 303). Although we are suggesting a different form of phase identification based on prevalent state sequences or distribution patterning, it remains an open question whether phase sequences could be useful performance differentiators. Further, based on our findings, entropy at the transition points was a predictor of performance, so perhaps the ordering of communication states at transition points matters more than the particular states comprising a phase, although this could vary as a function of team expertise.

Additional research could explore development of a near–real-time measure of phase transitions. For example, this could be based on the work of Gorman, Cooke, et al. (2010), which showed that perturbations to team communication could be detected in real time by calculating and updating estimates of the stability of the communication as the data are input from team members completing a task. Note that adopting this approach may require either the use of speaker-only codes (i.e., codes that only represent speaker and not semantic content; cf. Strang et al., 2012; Parker et al., 2016, when exploring alternative coding systems) or the use of methods from computational linguistics that can semantically categorize single utterances in near real time (e.g., Angus et al., 2012). Alternatively, one could use more fundamental linguistic aspects such as the categorization of vocal frequency (Black et al., 2013) or speech/pause dynamics (Fusaroli & Tylén, 2016).

Because conversations are comprised of multiple scales and levels with coupling between these (Abney, Paxton, Dale, & Kello, 2014), it is possible that phase transitions may have cascading effects across modalities, and as such the examination of many modalities and their crossover effects may be worthwhile (Barbosa, Déchaine, Vatikiotis-Bateson, & Yehia, 2012; Gorman et al., 2016). The long-term goal of such a real-time identification of phase transitions is the design of interactive systems that can administer timely prompts to enhance human cognitive performance alone and in teams (Wiltshire & Fiore, 2014; Wiltshire et al., 2014).

Another future direction is identifying the control parameters (Kelso, 2009) for CPS, which would involve answering this question: What parameters regarding the communication structures or the task, when varied, result in changes in the observed entropy values? As an example, Gorman et al. (2016) showed that higher level structures of their team training task were reflected in changes in entropy values representing team neurophysiological states. Similarly, Strang et al. (2015) showed that the coupling of team members' brain activity, measured with an entropy statistic, varied as a function of a task-load manipulation. While these are relatively macro-level variables, other control parameters may be microlevel features of a CPS task. Working toward identifying what these control parameters are, although task specific, will be informative in understanding CPS.

In terms of practical implications, we view the method applied here as versatile and easily adopted for other contexts. On the one hand, our technique can be easily applied to any type of categorically coded communications time series. This can include many other types of problem-solving such as tracking developmental transitions in children's problem-solving strategies (e.g., Siegler, 1994) or the solving of complex problems (Funke, 2010). But phase transitions seem pervasive across many types of social interaction (e.g., teams, couples, friends, coworkers; Butler, 2011) and, thus, our approach can apply to a variety of modalities and contexts. For example, our approach could apply equally well to many different behavioral modalities (see, e.g., Louwerse et al., 2012). In addition to laboratory tasks, it could be applied to other contexts such as education (Jordan & McDaniel, 2014), health care (Salas, Wilson, Murphy, King, & Salisbury, 2008), complex work teams in naturalistic settings (Fiore et al., 2014), therapy (Baucom et al., 2007), and others. The primary criteria that must be met in order to apply the sliding window entropy technique used here to detect phase transitions is that there are data in the form of discrete sequential time series. However, other data types could be assessed using other forms of entropy such as sample entropy or approximate entropy (Richman & Moorman, 2000). Concomitantly, theoretical advancements would be necessary to characterize the when and why transitions between phases occur in specific contexts as well as what those phases are.

5. Conclusion

In this paper, we integrated theory on CPS and dynamical systems theory regarding phase transitions. We also demonstrated a method for identifying phase transitions during CPS based on use of sliding window entropy and several criteria for identifying and verifying the robustness of peaks in entropy. From a theoretical perspective, this work contributes to dynamical systems theory by adding further validity to the claim that entropy can be used as an indicator of phase transitions. It also contributes to team cognition theory about CPS by providing empirical evidence that teams do exhibit phase transitions in their communications and that lower entropy at transition points is predictive of better CPS performance. These findings are consistent with related work on team performance and the more general theory of optimum variability. More broadly, this work enriches our scientific ability to understand CPS by integrating theories and providing a method to identify time points at which phase transitions occur. Further, we identified future directions that clearly build on our method for identification of phase transitions during CPS, and that will allow us to answer the question of "what" the phases exhibited by the team actually are and whether particular sequences of those phases can be related to CPS performance. We view the identification of phase transitions as foundational to understanding the coordinative structures of any system, whether that is a team collaboratively solving a problem in a high-stakes work domain or simply friends in an everyday interaction. Thus, the theory and methods presented here have far-reaching implications for research into human cognition and interaction.

Acknowledgments

The authors thank Cleotilde Gonzalez, Riccardo Fusaroli, and two anonymous reviewers whose feedback and comments on this manuscript were invaluable to improving the substantive contribution of our work. Also, we thank Aaron Likens and Sune Vork Steffensen for very helpful discussions regarding our methods. Thanks to Matthey J. Harvey for reading the manuscript and providing useful feedback that contributed to increased style and clarity of our writing.

Note

1. Measures of central tendency such as the mean, standard deviation, and range fail to capture the dynamics of team interaction (Gorman, Amazeen, et al., 2010; Gorman, Cooke, et al., 2010). Their meaning in the present context is even further reduced due to the discrete and nominal nature of the categorical communication sequences. Thus, Shannon entropy is the most appropriate choice because it preserves dynamics by reflecting the orderliness of the patterning inherent to a given sequence.

References

Abney, D. H., Paxton, A., Dale, R., & Kello, C. T. (2014). Complexity matching in dyadic conversation. *Journal of Experimental Psychology: General*, *143*(6), 2304.

Abney, D. H., Paxton, A., Dale, R., & Kello, C. T. (2015). Movement dynamics reflect a functional role for weak coupling and role structure in dyadic problem solving. *Cognitive Processing*, *16*(4), 325–332.

Amici, F., & Bietti, L. (2015). Coordination, collaboration and cooperation: Interdisciplinary perspectives. *Interaction Studies*, *16*(3), vii–xii.

Angus, D., Smith, A. E., & Wiles, J. (2012). Human communication as coupled time series: Quantifying multi-participant recurrence. *Audio, Speech, and Language Processing, IEEE Transactions on*, *20*(6), 1795–1807.

Araújo, D., Passos, P., Esteves, P., Duarte, R., Lopes, J., Hristovski, R., & Davids, K. (2015). The micro-macro link in understanding sport tactical behaviours: Integrating information and action at different levels of system analysis in sport. *Movement and Sport Sciences*, *89*, 53–63. doi:10.1051/sm/2015028

Bak, P., Tang, C., & Wiesenfeld, K. (1988). Self-organized criticality. *Physical Review A*, *38*(1), 364.

Bales, R. F., & Strodtbeck, F. L. (1951). Phases in group problem-solving. *The Journal of Abnormal and Social Psychology*, *46*(4), 485.

Banerjee, M., Capozzoli, M., McSweeney, L., & Sinha, D. (1999). Beyond kappa: A review of interrater agreement measures. *Canadian Journal of Statistics*, *27*(1), 3–23.

Barbosa, A. V., Déchaine, R. M., Vatikiotis-Bateson, E., & Yehia, H. C. (2012). Quantifying time-varying coordination of multimodal speech signals using correlation map analysis. *The Journal of the Acoustical Society of America*, *131*(3), 2162–2172.

Baucom, B., Eldridge, K., Jones, J., Sevier, M., Clements, M., Markman, H., Stanley, S., Sayers, S. L., Sher, T., & Christensen, A. (2007). Relative contributions of relationship distress and depression to communication patterns in couples. *Journal of Social and Clinical Psychology*, *26*(6), 689.

Bedwell, W. L., Wildman, J. L., DiazGranados, D., Salazar, M., Kramer, W. S., & Salas, E. (2012). Collaboration at work: An integrative multilevel conceptualization. *Human Resource Management Review*, *22*(2), 128–145.

Berg, C. A., Johnson, M. M., Meegan, S. P., & Strough, J. (2003). Collaborative problem-solving interactions in young and old married couples. *Discourse Processes*, *35*(1), 33–58.

Biron, H. C., Burkman, L. M., & Warner, N. (2008). A re-analysis of the collaborative knowledge transcripts from a noncombatant evacuation operation scenario: The next phase in the evolution of a team collaboration model (No. NAWCADPAX/TR-2008/43). Naval Air Warfare Center Aircraft Division, Patuxent River, MD.

Black, M. P., Katsamanis, A., Baucom, B. R., Lee, C. C., Lammert, A. C., Christensen, A., Gergiou, P. G., & Narayanan, S. S. (2013). Toward automating a human behavioral coding system for married couples' interactions using speech acoustic features. *Speech Communication*, *55*(1), 1–21.

Boker, S. M., & Laurenceau, J. P. (2006). Dynamical systems modeling: An application to the regulation of intimacy and disclosure in marriage. In T. A. Walls & J. L. Schafer (Eds.), *Models for intensive longitudinal data* (pp. 195–218). New York: Oxford University Press.

Bonachela, J. A., Hinrichsen, H., & Munoz, M. A. (2008). Entropy estimates of small data sets. *Journal of Physics A: Mathematical and Theoretical*, *41*(20), 202001.

Butler, E. A. (2011). Temporal interpersonal emotion systems the "TIES" that form relationships. *Personality and Social Psychology Review*, *15*(4), 367–393.

Butner, J., Amazeen, P. G., & Mulvey, G. M. (2005). Multilevel modeling of two cyclical processes: extending differential structural equation modeling to nonlinear coupled systems. *Psychological Methods*, *10*(2), 159.

Chen, G., Mathieu, J. E., & Bliese, P. D. (2004). A framework for conducting multilevel construct validation. *Research in Multilevel Issues: Multilevel Issues in Organizational Behavior and Processes*, *3*, 273–303.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*. Mahwah, NJ: Erlbaum.

Cooke, N. J., Gorman, J. C., Duran, J. L., & Taylor, A. R. (2007). Team cognition in experienced command-and-control teams. *Journal of Experimental Psychology: Applied*, *13*(3), 146.

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

Cooke, N. J., Gorman, J. C., & Rowe, L. J. (2004). *An ecological perspective on team cognition*. Mesa, AZ: Cognitive Engineering Research Institute.

Dakos, V., Van Nes, E. H., D'Odorico, P., & Scheffer, M. (2012). Robustness of variance and autocorrelation as indicators of critical slowing down. *Ecology*, *93*(2), 264–271.

Dale, R., Fusaroli, R., Duran, N., & Richardson, D. C. (2013). The self-organization of human interaction. *Psychology of Learning and Motivation*, *59*, 43–95.

Dale, R., Warlaumont, A. S., & Richardson, D. C. (2011). Nominal cross recurrence as a generalized lag sequential analysis for behavioral streams. *International Journal of Bifurcation and Chaos*, *21*(04), 1153–1161.

de Montjoye, Y. A., Stopczynski, A., Shmueli, E., Pentland, A., & Lehmann, S. (2014). The strength of the strongest ties in collaborative problem solving. *Scientific Reports*, *4*, 1–6.

Fiore, S. M., Rosen, M., Salas, E., Burke, S., & Jentsch, F. (2008). Processes in complex team problem solving: Parsing and defining the theoretical problem space. In M. Letsky, N. Warner, S. M. Fiore, & C. Smith (Eds.), *Macrocognition in teams* (pp. 143–163). London: Ashgate.

Fiore, S. M., Rosen, M. A., Smith-Jentsch, K. A., Salas, E., Letsky, M., & Warner, N. (2010). Toward an understanding of macrocognition in teams: Predicting processes in complex collaborative contexts. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *52*(2), 203–224.

Fiore, S. M., & Schooler, J. W. (2004). Process mapping and shared cognition: Teamwork and the development of shared problem models. In E. Salas & S. M. Fiore (Eds.), *Team cognition: Understanding the factors that drive process and performance* (pp. 133–152). Washington, DC: American Psychological Association.

Fiore, S. M., Smith-Jentsch, K. A., Salas, E., Warner, N., & Letsky, M. (2010). Towards an understanding of macrocognition in teams: Developing and defining complex collaborative processes and products. *Theoretical Issues in Ergonomics Science*, *11*(4), 250–271.

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: Cognitive Science*, *7*(1531), 1–17. doi:10.3389/fpsyg.2016.01531

Fiore, S. M., Wiltshire, T. J., Oglesby, J., O'Keefe, W., & Salas, E. (2014). Complex collaborative problem solving processes in mission control. *Aviation, Space, and Environmental Medicine*, *85*(4), 456–461.

Fischer, A., Greiff, S., & Funke, J. (2012). The process of solving complex problems. *Journal of Problem Solving*, *4*(1), 19–42.

Fischer, U., McDonnell, L., & Orasanu, J. (2007). Linguistic correlates of team performance: Toward a tool for monitoring team functioning during space missions. *Aviation, Space, and Environmental Medicine*, *78* (Supplement 1), B86–B95.

Fisher, B. A. (1970). Decision emergence: Phases in group decision-making. *Communications Monographs*, *37*(1), 53–66.

Fraser, A. M., & Swinney, H. L. (1986). Independent coordinates for strange attractors from mutual information. *Physical Review A*, *33*(2), 1134.

Funke, J. (2010). Complex problem solving: A case for complex cognition? *Cognitive Processing*, *11*(2), 133–142.

Fusaroli, R., Rączaszek-Leonardi, J., & Tylén, K. (2014). Dialog as interpersonal synergy. *New Ideas in Psychology*, *32*, 147–157.

Fusaroli, R., & Tylén, K. (2016). Investigating conversational dynamics: Interactive alignment, interpersonal synergy, and collective task performance. *Cognitive Science*, *40*(1), 145–171.

Gates, K. M., Gatzke-Kopp, L. M., Sandsten, M., & Blandon, A. Y. (2015). Estimating time-varying RSA to examine psychophysiological linkage of marital dyads. *Psychophysiology*, *52*(8), 1059–1065.

Gorman, J. C., Amazeen, P. G., & Cooke, N. J. (2010). Team coordination dynamics. *Nonlinear Dynamics, Psychology, and Life Sciences*, *14*(3), 265.

Gorman, J. C., Cooke, N. J., & Amazeen, P. G. (2010). Training adaptive teams. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *52*(2), 295–307.

Gorman, J. C., Cooke, N. J., Amazeen, P. G., & Fouse, S. (2012). Measuring patterns in team interaction sequences using a discrete recurrence approach. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *54*(4), 503–517.

Gorman, J. C., Hessler, E. E., Amazeen, P. G., Cooke, N. J., & Shope, S. M. (2012). Dynamical analysis in real time: detecting perturbations to team communication. *Ergonomics*, *55*(8), 825–839.

Gorman, J. C., Martin, M. J., Dunbar, T. A., Stevens, R. H., Galloway, T. L., Amazeen, P. G., & Likens, A. D. (2016). Cross-level effects between neurophysiology and communication during team training. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *58*(1), 181–199.

Guastello, S. J. (2000). Symbolic dynamic patterns of written exchanges: Hierarchical structures in an electronic problem solving group. *Nonlinear Dynamics, Psychology, and Life Sciences*, *4*(2), 169–187.

Guastello, S. J. (2011a). Entropy. In S. J. Guastello & R. A. Gregson (Eds.), *Nonlinear dynamical systems analysis for the behavioral sciences using real data* (pp. 213–230). Boca Raton, FL: CRC Press.

Guastello, S. (2011b). Orbital decomposition: Identification of dynamical patterns in categorical data. In S. J. Guastello & R. A. Gregson (Eds.), *Nonlinear dynamical systems analysis for the behavioral sciences using real data* (pp. 499–516). Boca Raton, FL: CRC Press.

Guetzkow, H. (1950). Unitizing and categorizing problems in coding quantitative data. *Journal of Clinical Psychology*, *6*, 47–58.

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: an overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, *8*(1), 23.

Harrison, S. J., & Stergiou, N. (2015). Complex adaptive behavior and dexterous action. *Nonlinear Dynamics, Psychology, and Life Sciences*, *19*(4), 345.

Hill, G. W. (1982). Group versus individual performance: Are N + 1 heads better than one? *Psychological Bulletin*, *91*(3), 517.

Hinrichsen, H. (2000). Non-equilibrium critical phenomena and phase transitions into absorbing states. *Advances in Physics*, *49*(7), 815–958.

Hirokawa, R. Y. (1983). Group communication and problem-solving effectiveness: An investigation of group phases. *Human Communication Research*, *9*(4), 291–305.

Hollenstein, T. (2007). State space grids: Analyzing dynamics across development. *International Journal of Behavioral Development*, *31*(4), 384–396.

Hollnagel, E. (2002). Cognition as control: A pragmatic approach to the modeling of joint cognitive systems. *Theoretical Issues in Ergonomic Science*, *2*(3), 309–315.

Hutchins, E. (1995). How a cockpit remembers its speeds. *Cognitive Science*, *19*(3), 265–288.

Hutchins, S. G., & Kendall, T. (2010). Analysis of team communications to understand cognitive processes used during team collaboration. In *Proceedings of the 15th International Command and Control Research and Technology Symposium (ICCRTS '10)*, Santa Monica, CA, June 22–24.

Jordan, M. E., & McDaniel Jr., R. R. (2014). Managing uncertainty during collaborative problem solving in elementary school teams: The role of peer influence in robotics engineering activity. *Journal of the Learning Sciences*, *23*(4), 490–536.

Kelso, J. A. S. (1990). Phase transitions: Foundations of behavior. *Synergetics of cognition* (pp. 249–268). Berlin: Springer.

Kelso, J. A. (2009). Coordination dynamics. In R. A. Myers (Ed.), *Encyclopedia of complexity and systems science* (pp. 1537–1565). New York: Springer.

Keyton, J., Beck, S. J., & Asbury, M. B. (2010). Macrocognition: A communication perspective. *Theoretical Issues in Ergonomics Science*, *11*(4), 272–286.

Klein, G., Ross, K. G., Moon, B. M., Klein, D. E., Hoffman, R. R., & Hollnagel, E. (2003). Macrocognition. *IEEE Intelligent Systems*, *18*(3), 81–85.

Kugler, P. N., & Turvey, M. T. (1987). *Information, natural law, and the self-assembly of rhythmic movement*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Levy, M. (2005). Social phase transitions. *Journal of Economic Behavior & Organization*, *57*(1), 71–87.

Likens, A. D., Amazeen, P. G., Stevens, R., Galloway, T., & Gorman, J. C. (2014). Neural signatures of team coordination are revealed by multifractal analysis. *Social Neuroscience*, *9*(3), 219–234.

Louwerse, M. M., Dale, R., Bard, E. G., & Jeuniaux, P. (2012). Behavior matching in multimodal communication is synchronized. *Cognitive Science*, *36*(8), 1404–1426.

Malloy, T. E., & Albright, L. (2001). Multiple and single interaction dyadic research designs: Conceptual and analytic issues. *Basic and Applied Social Psychology*, *23*(1), 1–19.

NASA (2011). Moonbase Alpha. Available at: http://www.nasa.gov/offices/education/programs/national/ltp/games/moonbasealpha/index.html. Accessed December 2, 2013.

Newell, A., & Simon, H. A. (1972). *Human problem solving* (Vol. 104, No. 9). Englewood Cliffs, NJ: Prentice-Hall.

OECD (2015). *PISA 2015 collaborative problem solving framework*. OECD Publishing. Available at: http://www.oecd.org/pisa/pisaproducts/Draft%20PISA%202015%20Collaborative%20Problem%20Solving%20Framework%20.pdf. Accessed December 8, 2015.

Palermos, S. O. (2016). The dynamics of group cognition. *Minds and Machines*, *26*, 409–440.

Parker, J., Best, C., Funke, G., Strang, A., & Marion, K. (2016). An investigation of coding schemes for sample entropy analysis of communications data. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 60, no. 1, pp. 111–115). Thousand Oaks, CA: SAGE Publications.

Pincus, D., & Guastello, S. J. (2005). Nonlinear dynamics and interpersonal correlates of verbal turn-taking patterns in a group therapy session. *Small Group Research*, *36*(6), 635–677.

Pincus, D., Ortega, D. L., & Metten, A. M. (2011). Orbital decomposition for the comparison of multiple categorical time-series. In S. J. Guastello & R. A. Gregson (Eds.), *Nonlinear dynamical systems analysis for the behavioral sciences using real data* (pp. 517–538). Boca Raton, FL: CRC Press.

Quesada, J., Kintsch, W., & Gomez, E. (2005). Complex problem-solving: A field in search of a definition? *Theoretical Issues in Ergonomics Science*, *6*(1), 5–33.

R Core Team (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at http://www.R-project.org/.

Ramseyer, F., & Tschacher, W. (2010). Nonverbal synchrony or random coincidence? How to tell the difference. In *Development of multimodal interfaces: Active listening and synchrony* (pp. 182–196). Berlin, Heidelberg: Springer.

Richardson, M. J., Dale, R., & Marsh, K. L. (2014). Complex dynamical systems in social and personality psychology. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (p. 253). New York.

Richman, J. S., & Moorman, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology-Heart and Circulatory Physiology*, *278*(6), H2039–H2049.

Rico, R., Sánchez-Manzanares, M., Gil, F., & Gibson, C. (2008). Team implicit coordination processes: A team knowledge–based approach. *Academy of Management Review*, *33*(1), 163–184.

Roschelle, J., & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem-solving. In C. E. O'Malley (Ed.), *Computer-supported collaborative learning* (pp. 69–97). Berlin: Springer-Verlag.

Rosen, M. A. (2010). Collaborative problem solving: The role of team knowledge building processes and external representations (Doctoral Dissertation). University of Central Florida, Orlando, FL.

Sadler, P., Ethier, N., Gunn, G. R., Duong, D., & Woody, E. (2009). Are we on the same wavelength? Interpersonal complementarity as shared cyclical patterns during interactions. *Journal of Personality and Social Psychology*, *97*(6), 1005.

Salas, E. E., & Fiore, S. M. (2004). *Team cognition: Understanding the factors that drive process and performance*. Washington, DC: American Psychological Association.

Salas, E., Wilson, K. A., Murphy, C. E., King, H., & Salisbury, M. (2008). Communicating, coordinating, and cooperating when lives depend on it: Tips for teamwork. *Joint Commission Journal on Quality and Patient Safety*, *34*(6), 333–341.

Schiepek, G., & Strunk, G. (2010). The identification of critical fluctuations and phase transitions in short term and coarse-grained time series—a method for the real-time monitoring of human change processes. *Biological Cybernetics*, *102*(3), 197–207.

Seeber, I., Maier, R., & Weber, B. (2013). Macrocognition in collaboration: Analyzing processes of team knowledge building with CoPrA. *Group Decision and Negotiation*, *22*(5), 915–942.

Shannon, C. E., & Weaver, W. (1959). *The mathematical theory of communication*. Urbana: University of Illinois Press.

Shavelson, R. J., & Webb, N. M. (1991). *Generalizability theory: A primer: A primer* (vol. 1). Thousand Oaks, CA: Sage Publications.

Siegler, R. S. (1994). Cognitive variability: A key to understanding cognitive development. *Current Directions in Psychological Science*, *3*(1), 1–5.

Srnka, K. J., & Koeszegi, S. T. (2007). From words to numbers: How to transform qualitative data into meaningful quantitative results. *Schmalenbach Business Review*, *59*(1), 29–57.

Steffensen, S. V., Vallée-Tourangeau, F., & Vallée-Tourangeau, G. (2016). Cognitive events in a problem-solving task: A qualitative method for investigating interactivity in the 17 Animals problem. *Journal of Cognitive Psychology*, *28*(1), 79–105.

Steiner, I. D. (1974). Whatever happened to the group in social psychology? *Journal of Experimental Social Psychology*, *10*(1), 94–108.

Stephen, D. G., Boncoddo, R. A., Magnuson, J. S., & Dixon, J. A. (2009). The dynamics of insight: Mathematical discovery as a phase transition. *Memory & Cognition*, *37*(8), 1132–1149.

Stephen, D. G., & Dixon, J. A. (2009). The self-organization of insight: Entropy and power laws in problem solving. *The Journal of Problem Solving*, *2*(1), 6.

Stephen, D. G., Dixon, J. A., & Isenhower, R. W. (2009). Dynamics of representational change: Entropy, action, and cognition. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(6), 1811.

Stergiou, N., & Decker, L. M. (2011). Human movement variability, nonlinear dynamics, and pathology: Is there a connection? *Human Movement Science*, *30*(5), 869–888.

Stevens, R. (2012). Charting neurodynamic eddies in the temporal flows of teamwork. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 56, No. 1, pp. 208–212). Thousand Oaks, CA: SAGE Publications.

Stevens, R. H., Galloway, T. L., Wang, P., & Berka, C. (2012). Cognitive neurophysiologic synchronies what can they contribute to the study of teamwork? *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *54*(4), 489–502.

Strang, A. J., Funke, G. J., Russell, S. M., Dukes, A. W., & Middendorf, M. S. (2014). Physio-behavioral coupling in a cooperative team task: Contributors and relations. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(1), 145.

Strang, A. J., Funke, G. J., Satterfield, K., Miller, B., Menke, L., & Brown, R. (2015). Effects of task-load transitions on EEG coupling in a high-tempo cooperative task verifying a basic utility for future team monitoring. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 59, No. 1, pp. 100–104). Thousand Oaks, CA: SAGE Publications.

Strang, A. J., Horwood, S., Best, C., Funke, G. J., Knott, B. A., & Russell, S. M. (2012). Examining temporal regularity in categorical team communication using sample entropy. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 56, No. 1, pp. 473–477). Thousand Oaks, CA: SAGE Publications.

Strough, J., Patrick, J. H., Swenson, L. M., Cheng, S., & Barnes, K. A. (2003). Collaborative everyday problem solving: Interpersonal relationships and problem dimensions. *The International Journal of Aging and Human Development*, *56*(1), 43–66.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson.

Tabachnick, B. G., Fidell, L. S., & Osterlind, S. J. (2001). *Using multivariate statistics*. Pearson Education.

Theiner, G., Allen, C., & Goldstone, R. L. (2010). Recognizing group cognition. *Cognitive Systems Research*, *11*(4), 378–395.

Weisberg, R. W. (2015). Toward an integrated theory of insight in problem solving. *Thinking & Reasoning*, *21*(1), 5–39.

Wildman, J. L., Salas, E., & Scott, C. P. (2014). Measuring cognition in teams a cross-domain review. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *56*(5), 911–941.

Wiltshire, T. J. (2015). Team interaction dynamics during collaborative problem solving (Doctoral Dissertation). University of Central Florida, Orlando, FL. http://digital.library.ucf.edu/cdm/ref/collection/ETD/id/6779

Wiltshire, T. J., & Fiore, S. M. (2014). Social cognitive and affective neuroscience in human-machine systems: A roadmap for improving training, human-robot interaction, and team performance. *Human-Machine Systems, IEEE Transactions on*, *44*(6), 779–787.

Wiltshire, T. J., Rosch, K., Fiorella, L., & Fiore, S. M. (2014). Training for collaborative problem solving: Improving team process and performance through metacognitive prompting. In *Proceedings of the Human Factors and Ergonomics Society 58th Annual Meeting* (pp. 1154–1158). Los Angeles, CA.

Appendix A: Communications coding (Wiltshire, 2015)

All dialog was time-stamped for onset of a given utterance, speaker demarcated, transcribed, unitized, and coded using a coding scheme based on the macrocognition in teams model (for similar methods, see Fiore et al., 2014; Rosen, 2010; Srnka & Koeszegi, 2007). The coding scheme, developed by Rosen (2010) to capture team knowledge-building processes, was used as the initial coding scheme to build upon and modify for the present research. Prior work successfully adapted and applied this scheme to a different problem-solving situation and, thus, demonstrated its generalizability (Fiore et al., 2014). The names of the codes and their brief definitions can be found in Table 1. A detailed codebook was developed specifically for this task that includes more detailed definitions as well as conditions for when and when not to use a code. This codebook was used by all coders.

Reliability was established for the unitization process using Guetzkow's U (Guestzkow, 1950), a test of the reliability of the number of units identified by independent coders, such that each unit only represented a single utterance or complete thought for which a single code could be assigned (Srnka & Koeszegi, 2007). This was calculated by comparing the numbers of utterances from each initial transcription to the coded transcriptions, in which coders were allowed to separate phrases within utterances in order to ensure that only a single code applied. The results showed high reliability for unitization, U = .002, with values closer to zero equaling high agreement.

Given that the codes are nominal and represent distinct collaboration processes, Cohen's kappa (κ) was determined to be the appropriate inter-rater reliability statistic to use (Hallgren, 2010). Further, because reliability of the coding was only established on a subset of all transcriptions (see below), κ was also selected given its relative conservativeness when compared to other inter-rater reliability statistics (Hallgren, 2010).

Two coders, who were blind to the hypotheses of the study, were trained to apply the coding scheme to three team's communication transcriptions. Training began by requiring the coders to familiarize themselves with the codebook. Then they were each assigned the same transcript and asked to go through and provide an initial code for each transcribed unit. Coders received feedback on their coding by providing them with a coded transcript that included all coders' assigned codes as well as coding done by the developer of the coding scheme. This allowed them to easily see where their coding was similar and different to the other coders. Each coder was then asked to provide a rationale for why their code may have been different from the others with specific guidance to refer to particular components of the codebook on which they based their coding. All rationales were compiled and distributed to all coders. The coders were then asked to determine a revised code for each utterance based on the coding comparisons and rationales. After this, coders were given feedback in terms of their percentage agreement with each of the other coders and a breakdown of codes that they were using inconsistently based on the percentages shown in contingency tables. This process was iterated three times for the training of the coders. Initial reliability of the coding during training was good ($\kappa = .67$) and by the completion of the training, reliability was excellent ($\kappa = .81$; Banerjee, Capozzoli, McSweeney, & Sinha, 1999).

Given the time intensity required for this detailed coding system, the two coders independently coded the same eight transcriptions (approximately one-fifth of the total remaining transcriptions) to establish inter-rater reliability. This is a greater proportion than the one-sixth recommended by Louwerse et al. (2012) for a more time-intensive coding process. Overall inter-rater reliability across these transcriptions was excellent (κ = .74) with a range of κ = .55 to .88 for individual transcriptions. Because reliability was excellent, the remainder of the data was evenly distributed and coded among the trained coders (16 transcripts each). To ensure coders remained reliable during the individual coding process, two random 40 unit excerpts were selected from uncoded transcriptions and assigned to each coder at the midpoint and completion of their individual coding, respectively. Results showed that throughout the individual coding, inter-rater reliability was excellent at the midpoint (κ = .81) and upon completion (κ = .71; cf. Banerjee et al., 1999).

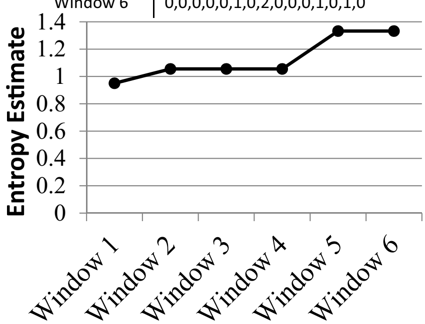
Appendix B: Illustration of sliding window entropy

On the right of Fig. 1B, there is a table representing a transcription excerpt detailing the unit number, the time, the utterance, the communications code, and the numerical

13,13,8,6,13,8,6,8,12,14,3,2,1,3,3,13,13,8,3,9,4,12,12,8,3,3,13,13,13,13,9,8,12,3,9,8,9,13,4,3,8,12,4,3,9

	Entropy
13,13,8,6,13	Window 1 = 0.9503
13,8,6,13,8	Window 2 = 1.0549
8,6,13,8,6	Window 3 = 1.0549
6,13,8,6,8	Window 4 = 1.0549
13,8,6,8,12	Window 5 = 1.3322
8,6,8,12,14	Window 6 = 1.3322

	Observed Code Counts
Window 1	0,0,0,0,0,1,0,1,0,0,0,0,3,0,0
Window 2	0,0,0,0,0,1,0,2,0,0,0,0,2,0,0
Window 3	0,0,0,0,0,2,0,2,0,0,0,0,1,0,0
Window 4	0,0,0,0,0,2,0,2,0,0,0,0,1,0,0
Window 5	0,0,0,0,0,1,0,2,0,0,0,1,1,0,0
Window 6	0,0,0,0,0,1,0,2,0,0,0,1,0,1,0



Unit #	Time-Min/Sec	Utterance/Action	Final Code	T12_Code
1	2:57	Okay.	INC/F/EX	13
2	2:58	Alright.	INC/F/EX	13
3	2:59	How do you want to start this off?	GTO	8
4	3:01	Uh, I say one of us repairs the, uh, solar panel and the other person repairs the the cable lines.	OG-P	6
5	3:07	Alright- let's see where we can we-	INC/F/EX	13
6	3:09	What do you want- What do you want to do?	GTO	8
7	3:12	I can repair the, uh, solar panels.	OG-P	6
8	3:14	Alrighty, I'll go get a holder and a wrench since I'm here.	GTO	8
9	3:17	Yeah.	S	12
10	3:36	Lovely weather we're having.	T/OT	14
11	3:42	Alright, come on...need a wrench.	KP	3
12	4:00	Hey, are there any black ones?	IR	2
13	4:01	Not that I can see.	IP	1
14	4:03	Which I guess is a good thing.	KP	3
15	4:06	Yeah, those we would have to replace.	KP	3
16	4:11	Wait, what do you mean I can't- what?	INC/F/EX	13
17	4:18	Okay.	INC/F/EX	13
18	4:32	Alright, so I guess we can begin by welding first instead of tightening.	GTO	8
19	4:36	Yeah, first we weld.	KP	3
20	5:24	Alright, I repaired one solar panel.	SU	9
21	5:26	Alright, do you have the, um, welder right now?	KR	4
22	5:28	Yeah.	S	12
23	5:29	Alright.	S	12
24	5:31	Do you need me to weld something?	GTO	8
25	5:32	Uh, yeah I need the, um, couplers.	KP	3
26	5:44	I'm not sure if I'm crazy, but it seems like we move aster if we jump.	KP	3
27	5:45	Which...	INC/F/EX	13
28	5:49	Maybe not that.	INC/F/EX	13
29	5:51	It might be true, in which case...	INC/F/EX	13
30	5:53	When do you....	INC/F/EX	13
31	5:56	Well, I got the wrench to check...uh, yeah.	SU	9
32	6:00	Alright, well I'll do this one that's right by the solar panel.	GTO	8
33	6:03	Alrighty.	S	12
34	6:10	I think jumping really does make you go faster.	KP	3
35	6:17	Alright, just, uh, secured that one.	SU	9
36	6:19	I'm gonna drop and pick up.	GTO	8
37	6:21	Secured that coupler.	SU	9
38	6:30	I'm gonna put the, um...	INC/F/EX	13
39	6:34	What else do we need over here?	KR	4
40	6:36	We're gonna need robots over there.	KP	3
41	6:39	For now I'm gonna put the wrench right here so it'll be close to where we need it.	GTO	8
42	6:43	Alright.	S	12
43	6:48	Where do we get those robots?	KR	4
44	6:54	I think I just set them down over there.	KP	3
45	7:05	We're saving some seconds.	SU	9

Fig. B1. Example translation from communication transcriptions to continuous entropy time series.

representation of the code. At the top of the figure a transposed version of the rightmost column of the transcription excerpt is shown. This represented the discrete nominal communications sequence. To illustrate the sliding window entropy, we show an example using a window size of five. The window includes the first five values, which are shown as observed counts below, and then entropy is calculated. Window 1 for example has an approximate entropy value of 0.95. Then the window is iterated by one unit to the right, adding the next consecutive value, and dropping the first value. We illustrate the sliding window using the first six windows in the time series and show the entropy value. These values are also shown in a line graph in the figure to show how the incremental addition and subtraction of values to the window captures changes in the distribution of values within a given window (see the observed code counts table) and in turn can be visualized through fluctuations in entropy.