Team Interaction Dynamics during Collaborative Problem Solving

2015

Travis Wiltshire

University of Central Florida

Find similar works at: http://stars.library.ucf.edu/etd

University of Central Florida Libraries http://library.ucf.edu

STARS Citation


This Doctoral Dissertation (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of STARS. For more information, please contact lee.dotson@ucf.edu.
TEAM INTERACTION DYNAMICS DURING COLLABORATIVE PROBLEM SOLVING

by

TRAVIS J. WILTSHIRE
B.S., University of Central Florida, 2009
M.S., Embry-Riddle Aeronautical University, 2012

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in Modeling and Simulation
in the College of Sciences
at the University of Central Florida
Orlando, Florida

Summer Term
2015

Major Professor: Stephen M. Fiore
ABSTRACT

This dissertation contributes an enhanced understanding of team cognition, in general, and collaborative problem solving (CPS), specifically, through an integration of methods that measure team interaction dynamics and knowledge building as it occurs during a complex CPS task. The need for better understanding CPS has risen in prominence as many organizations have increasingly worked to address complex problems requiring the combination of diverse sets of individual expertise to achieve solutions for novel problems. Towards this end, the present research drew from theoretical and empirical work on Macrocognition in Teams that describes the knowledge coordination arising from team communications during CPS. It built from this by incorporating the study of team interaction during complex collaborative cognition. Interaction between team members in such contexts has proven to be inherently dynamic and exhibiting nonlinear patterns not accounted for by extant research methods. To redress this gap, the present research drew from work in cognitive science designed to study social and team interaction as a nonlinear dynamical system. CPS was examined by studying knowledge building and interaction processes of 43 dyads working on NASA’s Moonbase Alpha simulation, a CPS task. Both non-verbal and verbal interaction dynamics were examined. Specifically, frame-differencing, an automated video analysis technique, was used to capture the bodily movements of participants and content coding was applied to the teams’ communications to characterize their CPS processes. A combination of linear (i.e., multiple regression, t-test, and time-lagged cross-correlation analysis), as well as nonlinear analytic techniques (i.e., recurrence quantification analysis; RQA) were applied. In terms of the predicted interaction dynamics, it was hypothesized that teams would exhibit synchronization in their bodily movements and complementarity in
their communications and further, that teams more strongly exhibiting these forms of coordination will produce better problem solving outcomes. Results showed that teams did exhibit a pattern of bodily movements that could be characterized as synchronized, but higher synchronization was not systematically related to performance. Further, results showed that teams did exhibit communicative interaction that was complementary, but this was not predictive of better problem solving performance. Several exploratory research questions were proposed as a way of refining the application of these techniques to the investigation of CPS. Results showed that semantic code-based communications time-series and %REC and ENTROPY recurrence-based measures were most sensitive to differences in performance. Overall, this dissertation adds to the scientific body of knowledge by advancing theory and empirical knowledge on the forms of verbal and non-verbal team interaction during CPS, but future work remains to be conducted to identify the relationship between interaction dynamics and CPS performance.
ACKNOWLEDGMENTS

A most sincere expression of gratitude goes to my academic advisor, Dr. Stephen M. Fiore. I cannot thank you enough for the professional training, academic opportunities, intellectual development and freedom, and continual support you provided as I completed my doctoral degree. Thank you for making this degree not just a possibility, but also a reality in which I could excel. In addition, I would like to thank Dr. Florian Jentsch, who I also viewed as an academic mentor. Your availability, guidance, attention to detail, and wealth of knowledge were pillars of my successes. I hope the amount of time each of you has invested in my academic development has already shown returns in the form of my scholarly productivity now and will continue to do so in the years to come.

Thank you to my committee members: Dr. Eduardo Salas and Dr. R. Paul Wiegand. Each of you has contributed insights and provided critical feedback that strengthened my research. Thank you for taking the time to support me and critically evaluate my research. To all of my committee members, your accomplishments and abilities as researchers set precedents that I aspire to achieve in my own career.

Many thanks go to the following colleagues: Rick Dale, Drew Abney, Alexandra Paxton, Riccardo Fusaroli, Moreno Coco, Adam Strang, and Jamie Gorman. Not only were your research and ideas fundamental to this research, but each of you was kind enough to lend me insights, feedback, and examples. Without your assistance, the methods and the ideas that I employed in this research would have been much more difficult to execute.
I would also like to acknowledge the many research assistants who were involved in this research: David Garcia, Sierra Snow, Michelle Sullivan, Amelaine Almacen, Omari Stringer, Savanna Hameroff, Neraj Ramnarace, Kelsie McCall, Hannah Nooshin, and Ashley Ercolino. Whether you contributed in a small way or a great way, you made this research possible and I appreciate the time and effort you dedicated.

I would like to thank the University of Central Florida, College of Graduate Studies for recognizing my accomplishments and awarding me with a Research Excellence fellowship that partially supported completion of this dissertation. And, I would like to acknowledge the Modeling and Simulation faculty Peter Kincaid and Sabrina Kalish for their pursuit of excellence and continual improvement of this program, which I am grateful to have been a part.

Lastly, I would like to express my gratitude for my family. In particular, thank you to my wife, Mercedes. You have always supported me in advancing my education and have helped me to find a great life balance. I look forward to our mutual support as we progress to the next chapter in our lives together. Mom and Dad: thanks for your support throughout the years.
# TABLE OF CONTENTS

LIST OF FIGURES .................................................................................................................. xi

LIST OF TABLES ...................................................................................................................... xii

CHAPTER ONE: INTRODUCTION .......................................................................................... 1

  Statement of the Problem ................................................................................................. 1

  Related Research and Deficiencies ............................................................................... 3

  Research Benefit ............................................................................................................. 6

  Purpose of the Research ............................................................................................... 7

CHAPTER TWO: LITERATURE REVIEW ............................................................................... 9

  Complex Problems ........................................................................................................ 9

  Collaborative Problem Solving .................................................................................... 11

  Interaction Dynamics .................................................................................................. 17

    Dynamical Systems Approach to Cognitive Science ................................................... 19

    Multiscale Coordination in Social and Team Interactions ............................................. 20

    Theoretical and Empirical Work on Interaction Dynamics ........................................... 23

    Interpersonal Synchronization ..................................................................................... 24

      Empirical Work Examining Interpersonal Synchronization ......................................... 28

    Complementarity of Interaction Sequences .................................................................. 35

  Specificity and Nonlinear Measures of Verbal Interaction Dynamics ......................... 40
CHAPTER THREE: METHOD ................................................................. 47

Design ................................................................................................. 47

Participants ......................................................................................... 48

Power Analyses and Sample Size ......................................................... 48

Materials ............................................................................................. 49

NASA Moonbase Alpha Simulation Task Description and Complexity ........................................... 50

Measures ............................................................................................. 54

Biographical Data ................................................................................ 54

Similarity Measure ............................................................................. 54

Performance Measures ...................................................................... 54

Interaction Measures .......................................................................... 55

Bodily Movement Time-Series ............................................................ 56

Communications Time-Series .............................................................. 57

Subjective Measures .......................................................................... 62

Social Value Orientation (SVO) ............................................................ 62

Cognitive Load Questionnaire ............................................................. 63

Knowledge Assessment ...................................................................... 64

Procedure ........................................................................................... 64
Communications Data Coding ................................................................. 66
Data Analyses ...................................................................................... 70
Frame-Differencing Technique ............................................................. 70
Recurrence Quantification Analysis ..................................................... 73

CHAPTER FOUR: RESULTS ..................................................................... 82
Data Screening .................................................................................... 82
Assessing Potentially Confounding Variables .................................... 82
Assessing Stationarity of Bodily Movement Time-Series ........................ 84
Testing Hypothesis 1: Emergence of Interpersonal Synchronization .......... 85
Testing Hypothesis 2: Does interpersonal synchronization predict team problem solving outcomes? ................................................................. 87
Testing Hypothesis 3: Complementarity in team communicative interaction sequences ........ 90
Testing Hypothesis 4: Does complementarity in team communications predict better problem solving outcomes? ......................................................... 97
Testing Hypothesis 5: Does higher synchronization in bodily movement and complementarity in team communications predict better problem solving outcomes, when compared to other groups? ................................................................. 99
Research Questions 1 & 2: Which communications time-series and which RQA-based measures are most sensitive to differences in problem solving performance? ....................... 100

CHAPTER FIVE: DISCUSSION ................................................................ 102
Theoretical and Practical Implications ........................................................................... 109

Limitations and Future Directions ............................................................................... 112

CHAPTER SIX: CONCLUSION ................................................................................. 121

APPENDIX A: LUNAR SURVIVAL MATERIALS ....................................................... 123

APPENDIX B: MOONBASE ALPHA TRAINING POWERPOINT .............................. 127

APPENDIX C: SOCIAL VALUE ORIENTATION ASSESSMENT ................................ 139

APPENDIX D: COGNITIVE LOAD RATING SCALE .................................................... 142

APPENDIX E: MOONBASE ALPHA KNOWLEDGE ASSESSMENT ............................ 144

APPENDIX F: CORRELATION MATRICES ................................................................. 147

APPENDIX G: COLLABORATIVE PROBLEM SOLVING CODEBOOK .................... 151

APPENDIX H: IRB APPROVAL DOCUMENT ............................................................ 182

REFERENCES ........................................................................................................... 184
LIST OF FIGURES

Figure 1. Macrocognition in Teams model (Fiore et al., 2014) ........................................................... 13
Figure 2. Screen shot representing profile view of participant recordings. ........................................... 50
Figure 3. Schematic representing the Moonbase Alpha settlement and flow of energy ...................... 51
Figure 4. Two bodily movement time-series for a dyadic team ............................................................ 57
Figure 5. Example team communications time-series ................................................................. 60
Figure 6. Example time lagged cross-correlation analysis plot ....................................................... 72
Figure 7. Three example recurrence plots generated from the speaker, code, and code*speaker time-series shown in Figure 5 ................................................................. 77
Figure 8. Example comparison of observed cross-correlation values versus surrogate cross-correlation values for one team ................................................................. 87
Figure 9. Analytic logic to test for complementarity interaction pattern ........................................ 92
LIST OF TABLES

Table 1  Definitions of team knowledge building processes (Adapted from Fiore et al., 2010b; Rosen, 2010) ......................................................................................................................... 14
Table 2  Varying definitions of interpersonal synchronization from extant literature. ........... 26
Table 3  Summary of research objectives, hypotheses, and questions. ................................. 45
Table 4  Potential confounding variables................................................................................ 46
Table 5  Team knowledge building codes to be used for communications data (adapted from Rosen, 2010). ............................................................................................................. 69
Table 6  Summary of RQA measures. ........................................................................................ 80
Table 7  Synchronization predicting performance multiple regression output ....................... 89
Table 8  Observed example of analytic logic illustrating evidence of complementarity interaction pattern. ....................................................................................................................... 94
Table 9  Communications transcription excerpt for high complementarity team. .................. 95
Table 10 Communications transcription excerpt for low complementarity team.................... 96
Table 11 Theoretical combinations of high vs. low synchronization and complementarity. .... 99
Table 12 Regression results from three communications time-series with RQA-based measures. .............................................................................................................................................. 101
CHAPTER ONE: INTRODUCTION

Statement of the Problem

Collaborative problem solving (CPS) is a specific area of team cognition research rising in prominence as public and private sector organizations are increasingly addressing complex problems requiring the combination of diverse sets of individual expertise to respond to novel situations (Fiore, 2008). Complex problems arise when there is a lack of knowledge for how to accomplish a goal, and determination of the solution requires the integration of knowledge across a large number of interconnected factors, often distributed across socio-technological systems (Fischer, Greiff, & Funke, 2012). Solving complex problems, thus, requires skilled teams that are able to adapt to high degrees of uncertainty, shifting task priorities, and dynamic systems and conditions; the effectiveness of which is driven by an amalgam of team members’ attitudes, behaviors, and cognition (Salas, Fiore, & Letsky, 2012). Unfortunately, without a full understanding of the ways effective teams collaborate during problem solving, teams are likely to fail to solve society’s increasingly complex problems.

A critical aspect of team cognition is the coordination of behaviors and knowledge in service of shared goals and objectives. This has traditionally been argued to be driven by shared and complementary knowledge structures across the team, as well as team member awareness of this distribution. Through concepts such as shared mental models and transactive memory systems, team cognition research has examined how the knowledge “held” by a team is related to performance (e.g., DeChurch & Mesmer-Magnus, 2010; Ren & Argote, 2011). Others have studied communication processes to infer similarities in knowledge held by a team that is related to process and performance (e.g., Bierhals, Schuster, Kohler, & Badke-Schaub, 2007). In the
study of CPS, specifically, extant team research has advocated for different facets of team cognition. Some research emphasizes team knowledge structures or the ways that teams build their knowledge in support of performance (e.g., Fiore et al., 2010a; Fiore, Smith-Jentsch, Salas, Warner, & Letsky, 2010b; Rico, Sánchez-Manzanares, Gil, & Gibson, 2008). Others have highlighted the need to understand team cognition as it emerges in the interaction of team members, often through their use of communicative behaviors (Cooke, Gorman, Myers, & Duran, 2013). Without proper integration of these two traditions, understanding how cognitive and collaborative factors dynamically unfold in service of CPS is fundamentally limited.

The critical issue of inquiry here is that CPS requires that a team engage in joint action in service of achieving their common goal of solving the complex problem at hand (e.g., Bedwell et al., 2012; Louwerse, Dale, Bard, & Jeuniaux, 2012). The process of collaboration, in service of problem solving, then centers on how team members, through their joint actions, build their knowledge through communicative interaction. In the study of collaborative joint action, emphasis is often placed on communicative dialog in which effective coordination has traditionally been argued to be characterized by the development of synchronized patterns of verbal and non-verbal behaviors (Baron, 2007), across varying temporal scales (Eiler, Kallen, Harrison, & Richardson, 2013), that ultimately lead to the alignment or shared understanding of the task or problem (Fiore & Schooler, 2004; Garrod & Pickering, 2009). However, a growing body of research is supporting the notion that the effective coordination of both verbal and non-verbal behaviors may rely on synchronous interaction processes as well as complementary interaction sequences that emerge during collaboration (e.g., Bernieri & Rosenthal, 1991; Cooke, et al., 2013; Coey, Vartlett, & Richardson, 2012; Sadler, Ethier, Gunn, Duong, & Woody, 2009;
Strang, Funke, Russell, Dukes, & Middendorf, 2014). This emerging body of work represents an important complement to traditional research on team cognition, which historically focuses on the measurement of static knowledge structures and only more recently, has emphasized interaction-based measures (e.g., Wildman, Salas, & Scott, 2014). However, because of the relatively recent emergence of interactive approaches to team cognition (Cooke et al., 2010), little is known with regard to these interaction dynamics in the context of CPS.

**Related Research and Deficiencies**

Recently, a model of collaborative problem solving was developed under the context of *Macrocognition in Teams* (MiTs; Fiore et al., 2010a, 2010b; Fiore, Wiltshire, Oglesby, O’Keefe, & Salas, 2014). The MiTs model focuses on explicating the knowledge-based collaborative and cognitive processes involved when teams work together to solve a unique and complex problem. Critically, the foundation for the model was a multi-theoretical and interdisciplinary integration uniting a number of disciplinary perspectives on team cognition. While the components and processes of the model are detailed in the literature review, the point here is that the MiTs model provides numerous predictions regarding the knowledge-based performance inherent to the collaborative problem solving process.

Given its recent emergence, work so far based on the MiTs model has been limited, yet promising. With the theoretical integration and prediction of CPS processes (Fiore et al., 2010a, b), most work so far has focused on identification of these processes in retrospective (e.g., Fiore et al., 2014) or simulated accounts of problem solving events with experienced domain practitioners (e.g., Hutchins & Kendall, 2010) or college populations (Rosen, 2010; Seeber, Maier, & Weber, 2013). One of the limitations of this body of work is its primary reliance on the
examination of verbal communications to understand CPS. Specifically, as argued above, the nature of joint action in CPS is both verbal and non-verbal, necessitating an understanding of how bodily movements and communicative acts relate to performance. Another limitation is that current analyses have either been descriptive or linearly predictive (e.g., Rosen, 2010), which fail to capture the dynamics of CPS.

To more fully test and refine the MiTs model, and inform team cognition more generally, better tools and methods are required to measure the processes associated with effective problem solving teams (cf. Salas, Cooke, & Rosen, 2008). Recent work has started to break ground in this area. For example, the Collaboration Process Analysis (CoPrA; Seeber et al., 2013) tool captures temporal aspects of team process during collaboration through use of an easy interface for coding and analysis of communications logs. In particular, Seeber et al. (2013) found support for Fiore et al.’s (2010b) team knowledge building processes during a CPS activity and described the inherent dynamics exhibited by differentially performing teams. But, CoPrA only captured these dynamics qualitatively and graphically. As such, it did not provide a means of quantifying the interaction dynamics of CPS or for examining the non-verbal behaviors associated with CPS. To build from this work, and redress this gap, this dissertation aims to address one of the limitations of team research in that it employs techniques that directly measure interaction dynamics. Importantly, these measures also help address another identified gap in the team cognition literature. Specifically, they are unobtrusive and capable of providing a fine level of temporal resolution to characterize the dynamics of CPS (e.g., Rosen, Bedwell, Wildman, Fritzsche, Salas, & Burke, 2011; Salas et al., 2008).
Teams and groups that engage in CPS are fundamentally a social-cognitive system (Larson & Christensen, 1993) that exhibits complex, nonlinear, and dynamic interactions (Richardson, Dale, & Marsh, 2014). As noted by Fiore et al. (2010a), CPS “can manifest in dynamic, iterative, recursive, and nonlinear ways” (p. 215). Put simply, the knowledge and behaviors of a team during CPS evolve and change over time. At issue is that interaction dynamics are not fully captured by traditional analytic techniques such as those employed by CoPrA or other research on MiTs. As such, the study of CPS teams should be treated as the study of a dynamical system (cf. Cooke et al., 2012) and draw on methods from complexity science (Fiore et al., 2010b).

Social interaction dynamics in humans are inherently multimodal (Louwerse et al., 2012) and display a variety of patterns indicating that individuals are coordinated through synchronization and complementarity (e.g., Sadler et al., 2009). For example, when engaging in a collaborative task, synchronization has been found to occur between humans in multiple modalities such as facial expression, gestures, linguistic communication (Louwerse et al., 2012; Dale & Spivey, 2006), eye-movement patterns (Dale, Kirkham, & Richardson, 2011), and inter- and intrapersonal limb movements (Richardson, Lopresti-Goodman, Mancini, Kay, & Schmidt, 2008).

However, recent research suggests that there are cases in which team members must engage in complementary, rather than synchronous, behaviors to effectively complete their task (e.g., Strang et al., 2014). For example, complementarity in the display of dominance behaviors was found in the form of an anti-phase relationship (i.e., when one person behaves dominantly the other does not and vice versa) between a dyad collaborating on a task (Sadler et al., 2009).
Further, Strang et al. (2014) found that complementarity in the physio-behavioral coupling of team members, who held differentiated roles, was predictive of task performance. More specifically, in Strang et al.’s (2014) study, higher degrees of synchronization in postural sway and cardiac interbeat intervals were negatively correlated with performance on a collaborative Tetris-based task.

Overall, the research on interaction dynamics is still nascent. As such, the role of the synchronous and complementary coordination of interaction behaviors, particularly as they relate to CPS, is relatively unknown. Given the emergence of synchronization and the facilitative coordinative structuring it provides to joint action (e.g., Knoblich et al., 2011; Lumsden, Miles, Richardson, Smith, & Macrae, 2012; Marsh, Richardson, & Schmidt, 2009), as well as the positive effect it has on rapport and cooperative predispositions (Hove et al., 2009; Valdesolo, Ouyang, & DeSteno, 2009; Wiltermuth & Heath, 2009), there is a critical need to better understand synchronous and complementary interaction dynamics in the context of CPS.

**Research Benefit**

The results of this experiment will have a number of theoretical and practical implications. Theoretically, the results aim to fill a critical gap in current understandings of the CPS processes that teams engage in, as well as determine the degree to which these empirical findings align with theoretical predictions. Further, this research can inform and advance scientific understanding of the role of synchronization and complementarity during team interaction dynamics and their relationship with effective collaborative problem solving. This would be of benefit to social, cognitive, and team scientists interested in topics such as joint action, coordination dynamics, team performance, and social cognition. In turn, this research
holds practical applications that, on the one hand, can be used to enhance current methods for researching CPS. For example, the research aims to provide new levels of specificity with regard to the types of verbal and non-verbal behaviors that should be examined, the various analytic measures that quantify their dynamics, and identification of those measures that are predictive of performance. On the other hand, the methods utilized here could be adopted as novel ways to monitor and enhance team performance in many complex work domains that require teams to collaboratively solve complex problems.

Purpose of the Research

In short, I united methods and concepts from a number of disciplines to study the roles of synchronization and complementarity as they relate to performance in teams engaged in a complex problem solving simulation. With regard to interaction dynamics, this work is informed by a recent theoretical account of dialog as interpersonal synergy proposing that synchronization may be essential for effective coordination at a lower, behavioral level (e.g., the non-verbal behaviors of a teammate) and, that at a higher, more linguistic level, complementary interactional routines are required (Fusaroli, Rączaszek-Leonardi, & Tylén, 2013). Therefore, in this research, I examined the emergence of interaction dynamics characteristic of synchrony and complementarity in the verbal and non-verbal exchanges between team members in terms of the relationship that these interaction dynamics have on effective collaborative problem solving performance.

Therefore, with the above as motivation for this work, the research objectives were to: (a) conduct an empirical investigation to better understand the verbal and non-verbal interaction dynamics of dyadic teams engaged in a simulated CPS task, (b) determine the relative effects of
certain patterns of interaction for a given interaction modality (i.e., bodily movement, team communications), and (c) to better understand the nature of the types of interaction dynamics that lead to better problem solving outcomes. Specifically, in this experiment, interaction dynamics were captured in two primary ways. Non-verbal behaviors of team members were captured at a gross level of granularity by using a frame-differencing technique that captures bodily synchronization in dyads (Paxton & Dale, 2013b). Verbal behaviors, representing the knowledge building processes inherent to CPS, were captured in communications transcriptions through use of content coding. These were then subject to nominal recurrence quantification analysis (e.g., Dale, Warlaumont, & Richardson, 2011; Gorman, Cooke, Amazeen, & Fouse, 2012a). In short, analytic techniques from dynamical systems theory (e.g., Richardson et al., 2014) and traditional linear analyses quantify, and are used to explain, the interaction dynamics with a specific focus on determining if these are predictive of effective CPS performance.
CHAPTER TWO: LITERATURE REVIEW

As a means of theoretically and empirically grounding this investigation on the nature of interaction dynamics during collaborative problem solving, a review of two major areas of research is required. First, complex problems are defined, and theoretical and empirical work on collaborative problem solving is reviewed. Next, interaction dynamics as they have been researched in the complementary, yet often distinct, studies of social and team interaction are detailed. An emphasis is placed on theories and supporting work that discuss synchronization of behavior during interaction and complementary interaction structures. Throughout this review of social and team interaction dynamics research, hypotheses for the present research are posited.

Complex Problems

As discussed previously, complex problems typically arise when there is a lack of knowledge for how to accomplish a goal, and determination of the solution requires the integration of knowledge from a large number of interconnected factors that may be distributed across varying socio-technological systems (Fischer et al., 2012). Ironically, increasing the complexity of the socio-technological systems required for operating in modern work domains such as aviation, aerospace, industrial process control, military, and even collaborative science, can give rise to equally complex problems (e.g., Letsky, Warner, Fiore, & Smith, 2008). Thus, this persistent technological evolution, in turn, motivates the need for a better understanding of complex problems and the cognitive and collaborative factors involved in solving them.

The key point in this section is not to cover in detail the theories and empirical work on complex problems as these can be found elsewhere (e.g., Fischer et al., 2012, Quesada, Kintsch,
& Gomez 2005; Wüstenberg, Greiff, & Funke, 2012). Instead, the focus here is simply to define what constitutes a complex problem to form the basis for discussion of collaborative problem solving.

Solving complex problems requires the adaptation to high degrees of uncertainty, shifting task priorities, and dynamic systems and conditions (e.g., Quesada et al., 2005). Because there is often a known goal, but a lack of knowledge for how to accomplish that goal, solving complex problems requires the generation of new knowledge (Wüstenberg et al., 2012; i.e., knowledge building; Fiore et al., 2010). The building of knowledge during complex problem solving is often a function of interacting with a system that has many interdependent elements and is thus complex because the elements of the system do not always relate to each other in a one-to-one manner (Quesada et al, 2005).

Complex problems are also dynamic and time-dependent, meaning that the problem solving environment can change at any moment, independent of the problem solver’s actions, and that certain actions need to be taken at certain points in time. These actions, in turn, affect future states of the system (Quesada et al., 2005). The very reasons articulated here are what differentiate complex problems from traditional, more basic, problems (cf. Hayes, 1989). In other words, problems, both basic and complex, are typically characterized by having a goal without a solution for how to reach that goal; however, problems increase in complexity as there are a larger number of factors to consider. In the Method section, the task that was selected for this research (i.e., NASA’s Moonbase Alpha simulation) qualifies as a complex problem in terms of the dimensions detailed here.
Extant research on complex problem solving has tended to focus on the cognitive processes of the individual. However, more than just individual knowledge building and system interaction is often required for complex problems. Solving many complex problems requires the collaborative efforts of teams who are able monitor and regulate their collective problem solving performance as they interact to integrate their perspectives and build their knowledge of the problem and potential solutions for solving it (Fiore et al., 2010a).

**Collaborative Problem Solving**

Collaborative problem solving (CPS) involves the coordinated, joint activity of multiple individuals as they adapt their existing knowledge or generate new knowledge to solve novel and complex problems (Fiore et al., 2010b). Indeed, research regarding how groups or teams solve problems together is not new (e.g., Bottger & Yetton, 1987; Hall, Mouton, & Blake, 1963); however, it has remained a relatively fragmented area of inquiry. Fiore and colleagues provided the theoretical integration and foundation necessary for gaining a more holistic understanding of collaborative problem solving processes (Fiore et al., 2010a) and predicted ways they may be observed and measured (Fiore et al., 2010b). This integration drew from theoretical and empirical work in areas such as computer-supported collaborative work, team cognition, group communication and problem solving theory, and work on distributed and extended cognition. The synthesis of these approaches are outlined in the Macrocognition in Teams model.

The theory of Macrocognition in Teams (MiTs) was developed to help understand team processes during complex CPS (Fiore et al., 2010a; 2010b) and more generally, as a way to characterize cognitive processes as they actually occur in real-world situations (e.g., Klein et al., 2003). MiTs focuses on ways in which internalized knowledge is transformed to externalized
knowledge by both individual and team cognitive processes. In other words, MiTs describes the ways that an individual may build their own internal knowledge and then as a collective, externalize and transform that knowledge such that it is made actionable by the team. In this way, MiTs addresses how teams collaboratively build knowledge in service of first gaining a collective understanding of the problem and then generating, evaluating, and executing effective problem solving solutions. MiTs acknowledges the importance of understanding how teams sequence their actions to accomplish a task, but focuses more on explicating the knowledge coordination of the team (Fiore et al., 2010b).

The MiTs model (Fiore et al., 2010a; 2010b) consists of five major components that characterize the collaborative problem solving process: individual and team knowledge building, internalized and externalized knowledge, and team problem solving outcomes (see Figure 1). *Individual knowledge building* occurs when an individual processes data and incorporates them 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-level 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 (or not openly challenged). *Team problem solving outcomes* are influenced by interactions among team members and whether these interactions contribute to fulfillment of critical task requirements (Fiore et al., 2010a). Teams with effective collaborative problem solving strategies engage in parallel and iterative processes that draw as necessary from
each of these components in service of constructing knowledge, understanding the problem, and evaluating potential solutions (Wiltshire, Rosch, Fiorella, & Fiore, 2014).

Figure 1. Macrocognition in Teams model (Fiore et al., 2014)

In terms of the MiTs model, the focus of this research will be primarily on those sub-processes of the team knowledge building component and the ways in which these affect the team’s problem solving outcomes. As articulated previously, team knowledge building involves the actions, primarily through communicative means, which team members exhibit to exchange and transform their collective knowledge into actionable ideas for solving the problem. The specific processes associated with team knowledge building are defined in Table 1 and are drawn from the extensive interdisciplinary integration of Fiore et al. (2010a; 2010b). These processes include team information exchange, team knowledge sharing, team solution option generation, team evaluation and negotiation of alternatives, and team process and plan regulation. To be an
effective CPS team, each of these processes, and the functions they play, should be incorporated into the team’s problem solving process. However, there is no necessary sequence dictating which order, and with what frequency, each process should constitute in the CPS process.

Table 1
*Definitions of team knowledge building processes (Adapted from Fiore et al., 2010b; Rosen, 2010)*

<table>
<thead>
<tr>
<th>Team Knowledge Building Process</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Team Information Exchange (TIE)</td>
<td>Exchanging relevant information with team members at the appropriate time</td>
</tr>
<tr>
<td>Team Knowledge Sharing (TKS)</td>
<td>Communicating explanations and interpretations of information with one or more team members</td>
</tr>
<tr>
<td>Team Solution Option Generation (TSOG)</td>
<td>Developing and offering potential solutions to a problem</td>
</tr>
<tr>
<td>Team Evaluation and Negotiation of Alternatives (TENA)</td>
<td>Clarifying and discussing positive and negative consequences of proposed solution options</td>
</tr>
<tr>
<td>Team Process and Plan Regulation (TPPR)</td>
<td>Critiquing the team’s process or plan given some further information or feedback about its potential for effectiveness</td>
</tr>
</tbody>
</table>

Now that an overview of the MiTs model and the specific components within that model that this research focuses on has been given, pertinent empirical work based upon MiTs is detailed next. Empirical research utilizing the Macrocognition in Teams model has begun to test components of this model individually, and also, identify evidence for the components collectively. This seems primarily to be a function of the type of data available to the researchers (e.g., team communications occurring while performing a problem solving task versus naturalistically-derived/field studies of problem-solving scenarios).

General support was found for the utility of the MiTs model by examining the communication logs, or transcriptions, of experienced teams performing tasks in domains such as North American Aerospace Defense command, Air Operations Center, and unmanned aerial
vehicle planning (Hutchins & Kendall, 2010). In examining data from across these domains, Hutchins and Kendall (2010) applied a categorical coding scheme based on the MiTs model with high reliability. Likewise, Hutchins (2011) found high reliability when applying a MiTs-based coding scheme to situation reports, forums, and blog posts included on a collaboration network for the Haiti disaster relief efforts. Evidence was also found for many of the CPS processes predicted by the MiTs model when examining retrospective accounts of a complex problem faced by diverse experts in NASA’s Mission Control Center (Fiore, Wiltshire, Oglesby, O’Keefe, & Salas, 2014). Notably, across all of these research projects, team knowledge building codes accounted for the greatest percentage of the data examined (at least 50% in all data sets). Indeed, the consistent robust support for team knowledge building processes across domains lends credence to its further emphasis in the present research.

With general support for the model provided by the naturalistic research above, laboratory and experimental work is pivotal for the present purposes. In a laboratory study, Rosen (2010) coded team communications data, derived from a simulated collaborative problem-solving task, to identify team knowledge building processes and determine their effect on problem solving performance. Indeed, this work forms the basis for how team communications will be analyzed in the present research (see Method section). Generally, the results showed that processes associated with team knowledge building were related to team problem solving outcomes. Specifically, those teams that shared more knowledge (i.e., higher frequency of team knowledge sharing) and less information had better problem solving performance. Further, this work also showed that the sequencing of team knowledge building processes was able to differentiate teams that performed well from those that did not. For example, high performing
teams showed a sequence of team knowledge building processes where each information exchange was acknowledged (i.e., closed loop communication); whereas, lower performing teams would only acknowledge after several information exchanges.

Other research found support for team knowledge building processes during a collaborative problem solving activity where teams were required to analyze and write a report regarding a fictitious information systems company (Seeber, Maier, & Weber, 2013). In particular, Seeber and colleagues (Frati & Seeber, 2013; Seeber et al., 2013) developed the Collaboration Process Analysis (CoPrA) tool to not only make communications coding easier, but to capture more of the temporal aspects of team knowledge building processes, at least qualitatively. That is, CoPrA provides visualizations to differentiate teams that “are problem-oriented/solution-minded, show consensus-oriented behavior, perform brainstorming by withholding criticism, discuss ideas in breadth and/or depth, or spend much effort on coordination” (Seeber et al., 2013, p. 939). However, although it is a useful tool for understanding temporal aspects of CPS, CoPrA only does so qualitatively through visual analysis of differences across teams.

Whereas the work based on MiTs to understand CPS is promising, much work remains. For example, Keyton, Beck, and Asbury (2010) highlighted a fundamental limitation of the MiTs model from a communications perspective: the development of meaning and evidence of cognitive activity by team members in their joint activity, both verbal and non-verbal. This approach is similar to interactive team cognition theory (Cooke, Gorman, Myers, & Duran, 2013) in that it views team cognition as an activity, that requires the team-level as the unit of analysis, and recognizes that team interaction is highly contextualized within a rich history of
prior knowledge and interaction. Indeed, the current research aims to evaluate both verbal and non-verbal activities of the team’s members as they collaborate to solve a complex problem while preserving the team as well as its interaction history in the analysis.

Another, more general limitation, is the analytic techniques employed to examine CPS in research based on the MiTs model. So far, these have remained primarily either descriptive, based upon qualitative data collected from retrospective reporting (e.g., Fiore et al., 2014), or linearly predictive (Rosen, 2010), based upon communication data collected during problem solving. However, as teams interact to solve a problem, their communications, verbal and non-verbal, are dynamic (i.e., changing over time) and multiscale (i.e., having different events occurring across a range of spatial and temporal scales). Although Seeber et al. (2013) provide a way to visualize the temporal unfolding of team knowledge building processes, they did not go so far as to quantify these dynamics. Likewise, Rosen (2010) noted that linear methods might be too simplistic to characterize the nature of CPS, as described by the MiTs model. To redress this gap and advance the science of collaborative problem solving, analytic techniques from dynamical systems approaches and complexity science should be adopted to investigate CPS (Fiore et al., 2010b). Such methods, as reviewed next, have been successfully applied to better characterize the nature of social and team interaction dynamics during a variety of tasks.

Interaction Dynamics

Collaborative problem solving is an inherently interactive process requiring the coordinated joint action of two or more individuals as they converse and act within the constraints of the task and in response to their interaction partners (e.g., Bedwell et al., 2012; Fiore et al., 2010a; 2010b; Keyton et al., 2010). The study of human interaction, generally, has a
rich and robust empirical history. It has continually progressed from studying social phenomena under an individualistic, static paradigm where participants make passive social judgments, to the more dynamic study of humans actually interacting with each other (cf. Baron & Boudreau, 1987; Good, 2007). For the present purposes, *actual social interaction* is characterized by the bidirectional and reciprocal engagement of two or more agents in which the mental states and behaviors of each agent, as well as environmental and situational features, dynamically influence the other (e.g., Bohl & van den Bos, 2012; Dale, Fusaroli, Duran, & Richardson, 2013; Przyrembel, Smallwood, Pauen, & Singer, 2012; Wiltshire, Lobato, Jentsch, & Fiore, 2013; Wiltshire, Lobato, McConnell, & Fiore, 2015).

Whereas social interaction can involve any sort of interaction amongst social agents fitting the general characteristics above, team interaction is a specific sub-type of social interaction characterized by the interaction partners being part of a team (e.g., Elias & Fiore, 2012). *Teams* are defined as “two or more people who interact dynamically, interdependently, and adaptively toward a common and valued goal/object/mission, who have each been assigned specific roles or functions to perform” (Salas, Dickinson, Converse, & Tannenbaum, 1992, p. 4). Because team interaction is a subset of social interaction, more generally, it is likely that research findings from social interaction have some general informative utility (e.g., Fiore, Salas, & Cannon-Bowers, 2001; Larson, & Christensen, 1993). Before discussing pertinent theoretical and empirical work examining social and team interaction dynamics, two key themes warrant brief discussion: description of the dynamical systems approach to cognitive science and multiscale coordination in social and team interaction.
Dynamical Systems Approach to Cognitive Science

With the present research, an emphasis on examining the dynamics of interaction aligns with the dynamical systems approach to cognitive science (e.g., Chemero, 2009; Dale et al., 2013; Marsh, Richardson, Baron, & Schmidt, 2006; Riley & Holden, 2012; Thelen & Smith, 1994; Vallacher & Nowak, 1997). A *dynamical system* is simply a system that changes over time as a function of interaction amongst the components comprising the said system (e.g., Richardson, Dale, Marsh, 2014). The dynamical systems approach is often described by contrasting it with the computational approach to cognition (e.g., Beer, 2000). Whereas the computational approach to cognition views cognition as performing logical operations on inputs (e.g., perceptions) to derive outputs (e.g., behaviors) “within” the individual (see Beer, 2000; Cooke, Gorman, Rowe, 2004) or even the team (e.g., Ilgen, Hollenbeck, Johnson, & Jundt, 2005), the dynamical systems approach treats the interaction of the system, the components comprising it, and the environment as the level of analysis, with a particular focus on the coordinative processes that characterize the behavior of the system over time (e.g., Coey, Varley, & Richardson, 2012). But, most importantly, dynamical systems approaches lend analytical techniques that are useful for understanding how a system changes over time, or how two systems interact and change together over time (Richardson et al., 2014).

In this way, the dynamical systems approach focuses more on the way the components of a system self-organize and mutually constrain each other to achieve some stable and coordinated performance (i.e., form), as opposed to examining evidence for how a system might be governed by a central-controller (e.g., executive functions; see for example Dale et al., 2013). Obviously, humans, in and of themselves, are complex systems with many interacting components and high
degrees of freedom, which only expand as the numbers of humans interacting increases. As such, another important concept from dynamical systems theory, that helps to address how a system with so many degrees of freedom can perform effectively, is that of a synergy: “a temporarily assembled, task-specific, functional coupling between a system’s componential degrees-of-freedom” (Coey et al., 2012, p. 4; see also Kelso, 2009). Synergies have been proposed as the means through which the degrees of freedom of complex dynamical systems are reduced and come together to form a functional unit in support of coordinated performance (e.g., Dale et al., 2013; Kelso, 2009).

Likewise, team cognitive processes, at the level of interacting individuals, have been described by Fiore and Salas (2004) in a way that is similar to that of a synergy in that the team’s cognitive processes fuse the various componential inputs of team members into a coherent functional whole. In particular, Fiore and Salas (2004) likened this to the binding problem in neuroscience where neural mechanisms act to fuse relevant sensory information into the phenomenology of consciousness. They note that coordinated team performance is analogously characterized by the synchronized actions of team members that are fused in service of effective processes. To better articulate and elaborate upon the degrees of freedom inherent to CPS teams and how they become coordinated in service of a task, multiscale coordination is described next.

**Multiscale Coordination in Social and Team Interactions**

Individuals and groups or teams of interacting individuals can all be construed as dynamical systems that exhibit multiscale coordination. This simply means that for a system to perform effectively in reaching some behavioral goal, it must be capable of coordinating its components across varying temporal and spatial scales (Eiler et al., 2013). Indeed, a range of
spatial scales, as small as molecular and neural events operating on the temporal order of less than milliseconds, up to societal and cultural scales operating across a temporal scale of years to generations, is argued to have some effect on coordination in social and team interactions (e.g., Fusaroli et al., 2014). At an individual level, the act of merely speaking requires the coordinated orchestration of the lips, tongue, jaw, lungs, and each of the relevant components of the nervous and musculoskeletal systems; all requiring precise spatial and temporal sequencing (Riley, Shockley, & Van Orden, 2012). During conversation in dyads, language itself exhibits multiscale coordination with temporal scales ranging from phonetic to lexical, semantic, and interpersonal, such as turn taking by interlocutors; all of which must be effectively coordinated to achieve an understanding of the conversational situation (Abney, Paxton, Kello, Dale, 2014). For the present purposes, the discussion of multiscale coordination is restricted to the forms of coordination that are involved in the non-verbal and verbal interactions inherent in teams who must collaborate to solve a complex problem.

From a joint action perspective, coordination may be either emergent or planned (Knoblich et al., 2011). *Joint action* is defined here as “a social interaction whereby two or more individuals coordinate their actions in space and time to bring about a change in the environment” (Knoblich et al., 2011, p. 60). In this context, *emergent coordination* encompasses the unintentional or spontaneous forms of coordination that occur during an interaction; whereas *planned coordination* encompasses those intentional aspects of interaction by which individuals plan their own actions, and, sometimes, the actions of others, towards accomplishing a given task goal. Indeed, these forms of coordination conceptually parallel the notions of implicit and explicit coordination described in research on team cognition (cf. Elias & Fiore, 2012; Rico et
al., 2012). While CPS is a form of joint action, the types of coordination briefly detailed here only describe the forms coordination may take and not what must be coordinated.

Importantly, with regard to team cognition and what becomes coordinated, Elias and Fiore (2012) noted that it is ultimately a “blend of internal mental activity and external social activity, which can be seen...as physical interactions between people” (p. 575). Given the necessity for teams to build their knowledge during CPS and the evidence of knowledge building during team communications (Fiore et al., 2010a, 2010b), drawing from psycholinguistic investigations of conversation during problem solving or collaborative tasks sheds light on what is coordinated during CPS.

At a general level, it is a commonly held view that through the interaction inherent to dialog or conversation, coordination of communicative behaviors is often in service of developing a shared situation model or common ground (e.g., Garrod & Pickering, 2004; Dale et al., 2014). This is a concept analogous to a shared problem model defined as a mental representation consisting of “situation- and task-appropriate strategies for interpreting and acting on a variety of task situations” (Fiore & Schooler, 2004, p. 128).

Whether such a representation exists or not, an understanding of the situation or problem must be formed. Prior work suggests that in order to do so, both content and process must be coordinated. Content coordination encompasses what is being said semantically by each interlocutor and thus involves coordination across each of the levels of individual and collective linguistic production during conversation (e.g., Delaherche et al., 2012; see also Pickering & Garrod, 2004 who call this form of coordination alignment). In contrast, process coordination involves the sequencing and timing of each individual’s actions, and the many degrees of
freedom associated with these actions, but also, importantly, the cognitive processes associated with the ability to predict an interaction partner’s behaviors. This allows, for example, individuals to take turns and have the ability to predict when an utterance might be completed by a conversation partner (Delaherche et al., 2012).

The discussion of coordination here is admittedly brief, though illustrative, at a high level, of the many components that require coordination during CPS. Interaction, in general, during CPS is complex and dynamic. Whether a team is collaboratively building its knowledge through discussion, proposing a solution for solving the problem, or performing some activity to execute a problem solution, team members’ verbal and non-verbal behaviors must be coordinated. However, the ways in which the interaction of team members is dynamically coordinated during CPS is not clear as this may often be a function of task and situation constraints (e.g., Strang et al., 2014). Therefore, theoretical and empirical work on social and team interaction dynamics are discussed next.

Theoretical and Empirical Work on Interaction Dynamics

As mentioned in the Introduction section, a major aim of this research is to better understand the nature of interaction dynamics during collaborative problem solving. In research on social and team interactions, theoretical and empirical work has developed which demonstrates that synchronized patterns of verbal and non-verbal behaviors (Baron, 2007; Fiore & Salas, 2004), across varying spatial and temporal scales (Eiler et al., 2013), ultimately lead to the alignment or shared understanding of the situation or problem and, in turn, effective coordination (e.g., Garrod & Pickering, 2009), or simply the fluid orchestration of a joint activity (Fusaroli et al., 2013). In an attempt to build on this work, other theoretical and empirical work
has posited and/or found that the effective coordination of both verbal and non-verbal behaviors may rely on synchronous as well as complementary interaction patterns (e.g., Bernieri & Rosenthal, 1991; Cooke, et al., 2013; Coey et al., 2012; Sadler et al., 2009; Strang, et al., 2014). Importantly, extant literature seems to lack a consistent definition of what constitutes synchronization during interaction. If synchronization and complementarity, in an interpersonal context, are two distinct patterns of interaction, then proper conceptualization and operationalization of these constructs is of critical importance to understanding CPS.

**Interpersonal Synchronization**

When it comes to social or team interaction dynamics, the meaning of synchronization is complex and difficult to define, given that there are divergent ways it is conceptualized and operationalized in the extant literature. Therefore, the focus in the first part of this section is to ensure proper clarification of terminology. Etymologically, the word *synchronous* originates “from the Greek words…(chronos, meaning time) and ...(syn, meaning the same, common), in a direct translation ‘synchronous’ means ‘sharing the common time’”, ‘occurring in the same time’. This term, as well as the related words ‘synchronization’ and ‘synchronized’, refers to a variety of phenomena in almost all branches of natural sciences, engineering and social life” (Pikovsky, Rosebblum, & Kurths, 2001, p. xvii). Further, Merriam-Webster (2014) defines the term *synchronize* as “to cause (things) to agree in time or to make (things) happen at the same time and speed.”

Synchronization was first discovered in a scientific context by Christiaan Huygens when he observed that pendulums on wall clocks became synchronized when they shared a common support. In particular, the oscillations of the pendulum were precisely in an anti-phase
synchronization pattern, meaning they were always moving in opposite directions, but temporally synchronized with regard to their oscillations (Pikovskiy et al., 2001). The idea of examining synchronization in human interaction was applied in a similar fashion to the discovery of synchronization of wall clock pendulums, only instead the focus was on inter-limb movements as participants swung hand-held pendulums (e.g., Schmidt, Beek, Treffner, & Turvey, 1991).

Much has progressed in the study of interpersonal coordination and interaction dynamics since the pioneering work in this area, particularly with regard to studying such phenomena in less artificial tasks; however, with progress, differential meanings associated with interpersonal synchronization have emerged. As a first attempt to parse this conceptual space, Table 2 provides an illustrative review of the varying definitions of synchrony.
Table 2  
*Varying definitions of interpersonal synchronization from extant literature.*

<table>
<thead>
<tr>
<th>Definition of Interpersonal Synchronization</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>“spatial and temporal behavior matching…a matching of behavior that is both time aligned like coordination and form aligned like imitation”</td>
<td>Louwerse et al. (2012) p. 2-3</td>
</tr>
<tr>
<td>“process by which two independent components continuously influence each other toward greater entrainment within a certain lag tolerance…or put more simply, to synchronize means that two entities through mutual influence come to do more or less the same thing within temporal proximity”</td>
<td>Fusaroli et al. (2013) p. 2</td>
</tr>
<tr>
<td>“the dynamic and reciprocal adaptation of the temporal structure of behaviors between interactive partners…the important element is the timing rather than the nature of the behaviors”</td>
<td>Delaherche et al. (2012) p. 3</td>
</tr>
<tr>
<td>“the degree to which the behaviors in an interaction are nonrandom, patterned, or synchronized in both timing and form…[where] interaction synchrony is composed of three components: rhythm, simultaneous movement, and the smooth meshing of interaction”</td>
<td>Bernieri &amp; Rosenthal (1991) p. 403</td>
</tr>
<tr>
<td>“correspondence between change in sound elements in the speech of a speaker and points of change in movement configurations shown by the listener”</td>
<td>Condon &amp; Sander (1974) p. 456</td>
</tr>
<tr>
<td>“the adjustment of the pace or cycle of an activity to match or synchronize with that of another activity. The adjustment could be in the phase, periodicity, magnitude, or some other temporal parameter of the activity. Pace refers to the speed at which an activity takes place. A cycle is a single complete execution of a periodically repeated phenomenon”</td>
<td>Ancona &amp; Chong (1999) p. 6</td>
</tr>
<tr>
<td>“similarity in rhythmic qualities and enmeshing or coordination of the behavioral patterns of both parties”</td>
<td>Burgoon, Stern, &amp; Dillman (1995) p. 128</td>
</tr>
<tr>
<td>“the way that interlocutors (individuals involved in conversation) grow to have similar behavior, cognition, and emotion over time”</td>
<td>Paxton &amp; Dale (2013b) p. 1</td>
</tr>
<tr>
<td>“Synchronized behaviors are those that are matched in time. Synchrony can occur with different actions, such as the coordinated movements of an athletic team or an orchestra; or with the same actions, such as pairs walking in stride”</td>
<td>Hove &amp; Risen (2009) p. 951</td>
</tr>
</tbody>
</table>

As should be evident from the etymological origins and these definitions, each description of synchronization emphasizes the temporal dimension; however, where the definitions diverge is on whether the spatial dimension (i.e., form of the behavior) must match as well in order to qualify the interaction of individuals as synchronized. Further, one definition
even includes an emotional dimension (Paxton & Dale, 2013). In all cases, the definitions are characterizing a form of interpersonal coordinated activity. On the one hand, the definitions of Bernieri and Rosenthal (1991), Louwerse et al. (2012), and Fusaroli et al. (2013) posit that synchronization is *both* the spatial and temporal matching of behavior. In other words, this suggests that if two people are synchronized during an interaction, they are displaying approximately the same behaviors at approximately the same time. Throughout the present work, the terms synchronized, synchrony, sync, and synchronous are used interchangeably to describe this phenomenon.

On the other hand, the definitions of Delaherche et al. (2012), Condon and Sander (1974), Ancona and Chong (1999), and Burgoon et al. (1995) all emphasize that synchronization stems from a reciprocity between interacting individuals and that the interaction is synchronous as long as both people are behaving to the same timing, rhythm, pace, or cycle. On this account of synchronization, the behaviors can be vastly different between interlocutors, but as long as they are neatly interwoven and appropriately timed, then this would constitute a synchronized interaction. In short, this characterization posits that synchrony is more a form of complementary coordination where the interaction is in service of some shared goal. Note that this latter definition is how synchronization is often conceptualized in team research (e.g., Ancona & Chong, 1999; Fiore & Salas, 2004). However, greater specificity can be gained with regard to understanding the interaction dynamics of collaborative problem solving when synchronization is defined as the former (i.e., spatial *and* temporal matching of behavior), as this allows for distinction between synchronization and complementarity and affords greater precision in measurement and analytical techniques (Paxton & Dale, 2013). However, to first strengthen this
empirical work examining synchronization across various modalities is reviewed in order to identify the degree to which the existing research aligns with the notion of synchronization as either including both spatial and temporal matching or primarily just including temporal matching. A focus is placed on research employing collaborative tasks. This review also serves the basis for discussion of complementarity in interaction and the empirical work associated with that phenomenon.

Empirical Work Examining Interpersonal Synchronization

Research has shown that synchronization is an emergent phenomenon that arises, often spontaneously, during various types of interaction and joint activity (e.g., Knoblich et al., 2011). A variety of contexts, as well as methods, have been used to examine synchronization. This includes detailed hand coding of videos of interaction frame by frame (Louwerse et al., 2012), trained judges rating the level of synchrony based on observation (Bernieri & Rosenthal, 1991), motion capture systems (Shockley, 2005), and automated video analysis (Paxton & Dale, 2013). Depending on the method selected, a variety of analytic techniques may also be employed including time-lagged cross-correlation, recurrence analysis, and spectral methods (Delaherche et al., 2012). Each method and analytic technique has their own methodological benefits and short-comings (see Bernieri & Rosenthal, 1991; Delaherche et al., 2012; Paxton & Dale, 2013, Richardson et al., 2014 for review). As appropriate, throughout this review, the method and analytical technique for examining synchronization is mentioned; however, those methods that are to be employed by the present research (i.e., time-lagged cross-correlation, recurrence quantification analysis) are described in detail in the Method section.
Results demonstrating the emergence of synchronized interaction dynamics are robust and can be found in a wide variety of settings. For example, when given instruction to oscillate their fingers up and down, participants with visual contact demonstrate an in-phase synchronization pattern, as measured with spectral analyses, that was significantly greater than those without visual contact (Ouiller De Guzman, Jantzen, Lagarde, & Kelso, 2008). Using spectral analysis of individuals sitting adjacent in rocking chairs, Richardson and colleagues (2007) found that participants intentionally and unintentionally synchronized their rocking pace and that the degree of synchrony increased when participants had visual information about each other (Richardson, Marsh, Isenhower, Goodman, & Schmidt, 2007). Using cross-recurrence quantification (a measure of the degree to which individuals exhibit similar behavioral patterns in time), in a study of staff-client discussion for individuals with intellectual disability, Reuzel et al. (2013) found that speech rhythms were synchronized. At a larger scale, in a study of soccer teams, individuals on the team, and even both teams, were shown to be synchronized in their overall bodily movements on the soccer field using cluster phase analysis (Duarte et al., 2013).

These are just a few of the examples that abound in the literature. The purpose here is to illustrate the range of relatively micro- (e.g., finger movements) to macro-scales (e.g., teams of teams) in which interpersonal synchrony has been observed to emerge. In accordance with the definitions of synchronization given by Bernieri and Rosenthal (1991), Louwerse et al. (2012), and Fusaroli et al. (2013), this collective body of research demonstrated that the alignment of both behavior and timing are required for an interaction to qualify as being synchronized. Next, several studies examining interpersonal synchronization in collaborative or cooperative tasks are reviewed with greater detail, as these are more relevant for the present study due to both the task
context and the method adopted for this dissertation (i.e., recurrence quantification, time-lagged cross-correlation).

In a study examining postural sway during a cooperative conversational task, Shockley, Santana, and Fowler (2003) found evidence of synchronization using cross-recurrence quantification analysis (CRQA). An in-depth discussion of recurrence quantification analysis is provided in the Method section. For the time being, it is important to note that this is a nonlinear measure for assessing the temporal dynamics characterizing the behavior of a system (e.g., Webber & Zbilut, 2005). It is often applied to examine the synchronization of two systems given its ability to easily convey whether or not two systems are exhibiting similar behavioral states at the same time (e.g., Delaherche et al., 2012; Fusaroli et al., 2014).

In Shockley et al.’s (2003) experiment, a dyad was tasked with cooperatively recognizing the differences between two similar images where each participant could see only one of the images. The experiment manipulated whether or not participants were actually able to see each other during the task as well as whether participants conversed with one another during the task or with a confederate. The conversations were qualified as cooperative when the two participants conversed because they were both trying to complete the same task, but not cooperative when participants conversed with the confederate because although the confederate discussed differences between the images, they were not vested in completing the task. CRQA was used to determine the percentage of recurrence between participants’ postural sway, which can be interpreted as the degree of synchronization in participants’ postural movements. Results showed that participants who actually conversed with each other during the task were significantly more
synchronized than those that conversed with a confederate. Further, being able to see the conversational partner did not have an effect in this study.

A similar task was employed by Dale, Kirkham, and Richardson (2011). Here dyads were each shown a set of tangrams (i.e., abstract shapes) on a screen, but the shapes were arranged in different orders. Thus, one participant was designated as the director and the other as the matcher and, through dialog, participants had to collaboratively determine a way to refer to the shapes so that the matcher could arrange the abstract shapes in the same way as the director. Dale et al. (2011) examined the degree to which the director’s and matcher’s eye movements were synchronized using CRQA. CRQA can measure synchronization between two systems by taking the diagonal recurrence profile, which indexes the degree to which the two systems are inhabiting the same state across different lags (e.g., Fusaroli et al., 2014). Synchronization is thus evidenced by a peak percentage of recurrence at lags that are close to 0 (see Method section for more details on this measure). Dale et al. (2011) found that, initially, the director’s eye movements were leading the matcher’s; however, across three rounds of completing this task, participants’ eye movements became increasingly synchronized with each other.

So far, the reviewed work only examines synchronization of a single behavioral modality (e.g., gaze direction, postural sway, etc.). Louwerse et al. (2012) investigated the degree to which multiple behavioral modalities become synchronized during a collaborative task. Like the tangram task, pairs of participants were each given a map with slightly different information and were assigned the roles of instructor or follower. The instructor was given a map with a route and it was their job to convey to the follower how to recreate that route on the map they were given, with inherent ambiguities between the two maps. Louwerse et al. (2012) captured, through an
intense and finely grained behavioral coding system, approximately 49 different behaviors of the face and head, gestures, and language. Using CRQA, as in the previously reviewed study, measures of synchronization were extracted for each behavioral modality, which were then compared to a baseline level of synchronization that could be expected due to chance. Results showed that, across all the major modalities, 19 of the behaviors demonstrated a pattern of synchronization between participants (e.g., laughing, nodding, acknowledgment, replies, descriptions, etc.). In other words, after a participant displayed a given behavior, for at least 10 of the behaviors, the other participant would display that same behavior in 5 seconds or less. Other behaviors had a slightly greater peak-lag time. Further, participants completed this task over repeated trials. Results showed that, as participants became more familiar with interacting with each other, synchrony increased in 63% of the modalities examined. Lastly, the difficulty of the task was also varied and results showed that synchronization was higher for 58% of the examined behaviors when the task was more difficult. These findings imply that synchronization is a robust and emergent phenomenon that occurs across many modalities, increases over time, and can be higher when tasks are more difficult as people interact and work together on collaborative tasks.

The fine-grained coding system applied by Louwerse et al. (2012) is very laborious to apply to a data set consisting of multiple dyadic interactions. In an attempt to partially overcome this challenge, Paxton and Dale (2013) developed a frame-differencing technique that provides a much more automated and less time intensive measure of overall bodily movement synchronization. In short, this technique compares each frame to the last and computes a standardized difference score that, with all else being held constant, represents a gross-level of
bodily movement for each participant. The next step of this technique involves conducting a time-lagged cross correlation analysis. When the peak cross-correlation is at a time lag of 0 and the value is significantly different from a baseline level of synchronization, this measure indicates that participants are synchronized in their bodily movements (see Method section for more elaboration on this analytic technique and its reliability; also see Figure 6). The only task in Paxton and Dale’s (2013b) experiment was for participants to discuss a topic of interest to both of them that was prompted by the researchers. The frame-differencing technique was applied to the videos to extract the degree of bodily synchronization for each dyad. These values were then used in several multiple regressions. Results showed that the degree of interpersonal synchronization was significantly reduced as time lag increased, lending support to the idea that the bodily movement of participants was synchronized during the conversation. Further, participants also rated the degree to which they liked each other. Results showed that these values moderated the degree of synchronization with those reporting higher liking also being more highly synchronized in their bodily movements.

Using the same frame-differencing technique, Paxton, Abney, Kello, and Dale (2014) investigated synchronization in a dyadic problem solving context. For this task, participants were required to build a tower together using only dry spaghetti and marshmallows. Each participant was only allowed to touch one of the resources, which required them to collaborate on who should use what resource and how. Like in prior work, Paxton et al. (2014) found evidence of synchronization in bodily movements during this problem solving task as well as in the speech onset/offset intervals of participants. Importantly, improved performance on the task, measured
in terms of a ratio of height/weight of the tower, was predicted by synchronization in bodily movement, as well as speech.

In addition to better task performance, examples of the observed benefits of synchronization in other contexts are as follows. Hove and Risen (2009) found that inducing interpersonal synchrony, by requesting participants to tap their fingers in sync, increased reported ratings of affiliation between participants. Valdesolo et al. (2010) induced interpersonal synchrony by having participants rock in a rocking chair in sync with one another and found that this improved perceptual sensitivity on a subsequent joint action task. Cirelli, Einarson, and Trainor (2014) induced synchrony between infants and caregivers and found that this increased the degree to which the infant will engage in a subsequent prosocial helping behavior. Conversely, Marsh et al. (2013) found that children with autism were not able to synchronize with their parent while rocking in a rocking chair; whereas, non-autistic children did. Ramseyer and Tschacher (2011) found that nonverbal synchrony in bodily movements between patient and therapist during a psychotherapy session was predictive of relationship quality and therapy outcome. Wiltermuth and Heath (2009) found that participants who walked in synchrony were more likely to engage in cooperative behaviors during a later economic exchange game. Lastly, Lumsden et al. (2012) found that participants with a cooperative predisposition were more likely to synchronize with others.

Taken together, these studies demonstrate that it is primarily both the form and timing of behaviors that must coincide to constitute interpersonal synchronization during an interaction. The findings also provide general evidence for its widespread emergence in a variety of tasks. However, the function of this phenomenon, particularly, in the context of complex collaborative
problem solving, is not well known. Nonetheless, given what appear to be robust findings of interpersonal synchronization across a variety of tasks and modalities, it is quite likely that synchronization will emerge during the CPS task employed by the present experiment. Further, given what appears to also be robust benefits for this form of interaction, it is also likely that synchrony will contribute a performance benefit during the problem solving task. Therefore, I hypothesize:

**Hypothesis 1:** Interpersonal synchronization, in the bodily movements of participants, will emerge during dyadic collaborative problem solving beyond a baseline level of synchrony due to chance.

**Hypothesis 2:** Higher degrees of interpersonal synchronization in the bodily movements of participants will be predictive of better team problem solving outcomes.

The goal of this brief review was to justify the operationalization of interpersonal synchronization adopted for this research by drawing on the etymological, conceptual, and empirical work related to this concept. As mentioned previously, interpersonal synchronization, alone, is unlikely to be able to account for the complexity of the interaction dynamics inherent to collaborative problem solving teams, especially with regard to characterizing team knowledge building communications. Therefore, in the section that follows, theory and research related to complementary interaction dynamics are discussed.

**Complementarity of Interaction Sequences**

A recent theoretical account, characterizing dialog as an interpersonal synergy (Fusaroli et al., 2013), and some related empirical support, has proposed that synchronization may be
essential for good coordination at a lower behavioral level (e.g., the non-verbal behaviors of a teammate) and conversely, that at a higher, more linguistic level (e.g., team knowledge building communications), complementary interactional routines are required (see Dale et al. 2013; Fusaroli et al., 2013). The major driving force behind this theoretical perspective is that, often in social and team interaction, individuals comprising these systems must say very different things to perform effectively, as opposed to indiscriminately saying the same things (e.g., Vatikiotis-Bateson, 2008). As defined previously, synergies characterize how, through context and task constraints, interaction partners become coupled and their degrees of freedom are reduced such that the system of individuals performing the task becomes a functional unit in a way that best supports meeting task demands (Coey et al., 2012; Kelso, 2009). Indeed, the idea of synergies or synergistic performance is not new to team research (see Stagl, Burke, Salas, & Pierce, 2006; Stasio, 2010; Tesluk, Mathieu, Zaccaro, & Marks, 1997); however, the tools and techniques of dynamical systems approaches that often ensue with such conceptualizations are much less widely adopted (see, however, Gorman et al., 2010; Gorman et al., 2012a; Russell et al., 2010; Strang et al., 2014). Regardless, the point here is that the coupling of interacting components (e.g., individuals and their parts across multiple spatial and temporal scales), form task-specific synergies, which may give rise to synchronous interaction dynamics as well as complementary interaction dynamics (Dale et al., 2013).

What exactly does complementarity mean in the sense of interaction dynamics? Fusaroli et al. (2013) defined it “as the way components doing quite different things come to form a coherent whole. A structure initiated by one component is completed (rather than copied) by the other” (p. 6). Likewise, Strang et al. (2014) described complementary behavior, in the context of
team performance, as overriding the emergent tendency to match the behaviors of a team member and engaging in opposite or asynchronous actions. To make this a bit more concrete, some of the most basic examples of complementarity, when it comes to dialog, take the form of adjacency pairs (i.e., a conversation unit comprised by two utterances by two speakers where the first utterance prompts a certain type of response utterance; Fusaroli et al., 2013; Mills, 2014).

Examples of general complementary actions or conversational adjacency pairs include: when one person is speaking, the other is listening; when one person asks a question, the other follows with a response; or when one person takes a step forward while moving a large piece of furniture, the other person takes a step back. In this way, the behaviors are complementary and are formed from the dynamical and mutual adaptation characterizing the interaction (Sadler et al., 2009).

In the context of complex collaborative problem solving, the collaborative processes conveyed through team communications necessitate this complementarity. For example, if one team member requests information, then another should provide it; likewise, if one team member generates a potential problem solution, then another should evaluate it (see Table 5 for detailed operationalizations of the team knowledge building communications codes). Empirical work examining complementarity, in this sense, is less robust than that of synchronization; however, several key examples are reviewed here.

Most research on complementarity of an interaction has examined interpersonal complementarity, which, rather than the dialogical adjacency pairs described above, is more concerned with how some higher-level category of interpersonal behavior (e.g., dominance) displayed during an interaction complements some other category displayed by another (e.g., Markey, Funder, & Ozer, 2003). In an experiment employing a collaborative task in which two
participants were given a set of stories and had to ultimately come to a collective decision regarding the personality of the person who developed the stories, interpersonal complementarity was investigated using a novel technique (Sadler et al., 2009). Specifically, the novel technique used was a post-hoc, dynamic rating technique where trained individuals rated videos of the interaction using a joystick to indicate dimensions of the interpersonal traits of affiliation and dominance during dyadic problem solving. The results showed, using spectral analytic techniques, that affiliation between interaction partners was synchronized, but dominance was complementary. In other words, when one person was high in affiliation, so was the other; but conversely, when one person demonstrated high dominance in the interaction, the other displayed low dominance (i.e., submissiveness).

Examples like this abound in the social and personality psychology literature with research showing, for example, that when complementary interpersonal behaviors are displayed, participants will be more satisfied with their interaction and view their interaction partners more similarly (Dryer & Horowitz, 1997; Tiedens & Fragale, 2003). Further, on an act-by-act basis, a single interpersonal behavior can predict a subsequent, complementary behavior by the interaction partner (Tracey, 1994; see also Sadler et al., 2011). However, what is of greater interest for the present research is the type of complementarity initially discussed in terms of conversational acts that are in service of collaborative problem solving.

Although little research has been done on the precise sequencing of team knowledge building communications during CPS (see, however, Rosen, 2010), the Macro-cognition in Teams model (MiTs) predicts that team members will move through problem solving phases and this is evidenced by the type of communications exchanged (e.g., Fiore et al., 2010a; 2010b).
The proposed research does not examine evidence for such phases, but the general theoretical motivation that can be derived from this is that, in order to reach an effective and efficient problem solving outcome, an iterative, yet dynamic and importantly, complementary sequencing of team knowledge building communications is necessary. For example, as mentioned above, the MiTs model prescribes that teams will iterate through the various team knowledge building processes to request and provide information and knowledge, to generate solutions to the problem and evaluate them, and to both assess and regulate their execution of the planned solution (Fiore et al., 2010a, 2010b). Through interaction, team members come together, each with their unique contributions, and form a coherent synergistic problem-solving unit.

Ultimately, what this interactive exchange during CPS amounts to, drawing from the more general study of conversation during joint problem solving, is a continual identification of the next relevant contribution (Clark, 1996; Mills, 2014). Indeed, with adjacency pairs, the next relevant contribution is straightforward (e.g., a request for information is followed by a provision of information, or a request for clarification), but it is not always clear to interlocutors what the appropriate contribution might be. The factors that contribute to the action selection by team members during CPS are the many task-constraints, context cues, and coupling of interlocutors, which, across spatial and temporal scales, must become coordinated to function effectively.

In this vein, I predict that complementarity more strongly characterizes the communicative interaction dynamics of teams engaged in CPS when compared to synchronization. The justification for this is that complementarity would be a more accurate depiction of the iteration through relevant team knowledge building processes and problem solving phases (see Fiore et al., 2010a, 2010b). Importantly, this research aims to describe the
interaction dynamics inherent to CPS teams at multiple scales and thus, drawing from the above theoretical and empirical work, the following hypotheses are posited:

**Hypothesis 3:** Teams will exhibit knowledge building communications that are more characteristic of a complementarity interaction pattern than a synchronized interaction pattern.

**Hypothesis 4:** Higher degrees of complementarity in team knowledge building communications will be predictive of better team problem solving outcomes.

**Hypothesis 5:** Teams that exhibit both high synchronization in their non-verbal behavior (i.e., bodily movements) and high complementarity in their team knowledge building communications will have better team problem solving outcomes when compared to other combinations of these metrics.

**Specificity and Nonlinear Measures of Verbal Interaction Dynamics**

While the above literature review has focused primarily on research examining interpersonal synchronization and complementarity, which, admittedly, are the primary focus of this research, other relevant work in team interaction dynamics can also prove insightful for the present investigation. Recall that interaction in a social or team context is inherently dynamic and exhibits nonlinear patterns (e.g., Richardson et al., 2014; Gorman et al., 2012a; Fiore et al., 2010a, 2010b). Hence, given that teams constitute a complex dynamical systems, nonlinear analytic techniques (e.g., recurrence quantification analysis) are well-suited for gaining insights into the way these systems behave spatially and temporally. Before progressing to discuss the specific method for the proposed research, two items require discussion. The first is the level of
specificity with which team communications should be captured. The next is that recurrence quantification analysis provides a number of measures regarding the dynamics of a system that can then be subject to more traditional, linear analytic techniques (e.g., Zbilut & Webber, 2006).

One of the important aspects of this research is the right level of specificity with regard to examining verbal interaction dynamics. For example, some research has examined only which team member was speaking during an interaction, represented as a single discrete sequential time-series (e.g., Gorman et al., 2012a). Others have examined the semantic content of an utterance conveyed by a team member where each speaker is relegated to their own separate time-series (e.g., Louwerse et al., 2012). Others, still, have sought to examine time-series that account for speaker, semantic content regardless of speaker, and semantic content that is speaker specific (e.g., Bowers, Jentsch, Salas, & Braun, 1998; Russell et al., 2012). Only a handful of authors have published work in this domain; therefore, the literature does not provide clear guidelines for which type of communications time-series to employ. Often this can be a function of how many people are involved in the social or team interaction. For example, if there are only two people, then cross-recurrence quantification (CRQA) can easily be used; however, if more people are involved, then it is not suitable to use CRQA and, thus, traditional RQA should be used instead (see Method section for the distinction between these two techniques). Further, the candidate communications time-series to use also depends on the specific question being employed. For example, if the research question is aimed at identifying synchronous patterns of communication between two individuals, then CRQA would be appropriate. But, if the goal is to identify complementarity, then RQA, with the entire communications as a single time-series, is
likely most sensitive (Fusaroli et al., 2014). Therefore, the following exploratory research question is posited:

**Research Question 1:** Which specificity of team communications time-series (e.g., speaker, semantic content, or speaker by content) provides the most insight and sensitivity to differences in problem solving outcomes?

As was partially alluded to in the prior literature review, but which is also discussed in detail in the Method section, recurrence quantification analysis (RQA) provides a number of measures that can quantify different aspects of the dynamics of verbal interactions (or any other system under study). In addition to just quantifying how much recurrence (i.e., how many times a certain state is revisited by the system), other measures from RQA can indicate how deterministic (i.e., rigid vs. flexible), how complex/regular (entropy), and how stable a system is (meanline) over the course of an interaction. While the precise details for each of these measures are provided in the Method section (see also Table 7), a brief review of some of this work follows.

In a study investigating determinism of team communication structures, representing solely who was speaking, Gorman et al. (2012a) investigated differences in team performance across a series of unmanned vehicle tasks, while varying whether a team was kept intact or mixed. Essentially, determinism, in this context, was an RQA derived metric that quantifies the temporal regularity of a time-series and indexes whether there are interaction patterns that repeat often in the data (high determinism) or interaction patterns that do not repeat as often (low determinism). Using RQA, the results showed that teams kept intact demonstrated significant increases in determinism in their communication patterns as they performed more tasks together.
Conversely, teams that were mixed across tasks did not show differences in determinism. This was interpreted as an index, in which teams kept intact formed more rigid communication patterns; whereas, mixed teams demonstrated more flexible communication patterns.

In research investigating the determinism of team communications during an air battle management task, Russell et al. (2012) varied whether a team was cross-trained or not, high versus low task demand, the type of collaboration technology teams were able to use to complete the task (i.e., radio-only, or both radio and text), and the way resources were displayed (tabular, graphical). As mentioned previously, Russell et al. (2012) examined three different communications time-series representing either who was speaking, the general semantic content, or the semantic content linked to who was speaking. Using RQA to derive the measures of determinism for each of the time-series, results showed no effects for the time-series indicating solely who was speaking. Note this is in contrast to work by Gorman and colleagues (2012). However, determinism was found to be higher when participants used the graphical versus tabular resource displays when examining the semantic content time-series. Further, the combined time-series yielded a number of interaction effects suggesting that when a team was cross-trained and had access to both collaboration technologies, their communications exhibited lower determinism. While Russell et al. (2012) interpreted this as a more chaotic communication structure, it could also be interpreted as more flexible and thus indicating those teams were more adaptable (Gorman et al., 2012a).

Although only a handful of experiments such as these exist, the point here is that RQA can provide additional insights by investigating differences in the various measures that quantify the dynamics of the team communication or the general system under study. Full elaboration of
the measures of interest and their interpretation in the context of team communications are provided in the Method section. Therefore, as a general research question, the following is proposed:

**Research Question 2:** Can measures derived from RQA, such as recurrence, determinism, meanline, and entropy, of different communication time-series be used to predict problem solving performance?

**Research Summary Information**

In sum, the aim of this literature review was to provide the reader with an understanding of complex problems and current theoretical and empirical work for how teams collaboratively solve such problems. An emphasis was placed on developing a better theoretical and empirical understanding of the interaction dynamics inherent to this process through elaboration on interpersonal synchronization and complementarity. As a means of refreshing the reader of the research objectives, hypotheses, and research questions, all of these are reiterated in Table 3. Lastly, potential confounding variables for the proposed research are detailed in Table 4.
Table 3
Summary of research objectives, hypotheses, and questions.

<table>
<thead>
<tr>
<th>Research Objectives</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Objective 1</strong></td>
</tr>
<tr>
<td><strong>Objective 2</strong></td>
</tr>
<tr>
<td><strong>Objective 3</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Research Hypotheses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hypothesis 1</strong></td>
</tr>
<tr>
<td><strong>Hypothesis 2</strong></td>
</tr>
<tr>
<td><strong>Hypothesis 3</strong></td>
</tr>
<tr>
<td><strong>Hypothesis 4</strong></td>
</tr>
<tr>
<td><strong>Hypothesis 5</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Research Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Research Question 1</strong></td>
</tr>
<tr>
<td><strong>Research Question 2</strong></td>
</tr>
</tbody>
</table>
Table 4  
*Potential confounding variables.*

<table>
<thead>
<tr>
<th>Name</th>
<th>Affected Variables</th>
<th>Method for control</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Video game experience</td>
<td>Task Performance</td>
<td>Covary</td>
<td>Orvis, Horn, &amp; Belanich (2008)</td>
</tr>
<tr>
<td>Familiarity with interaction partner</td>
<td>Bodily Synchronization</td>
<td>Control</td>
<td>Brennan &amp; Hanna (2009)</td>
</tr>
<tr>
<td>Interpersonal similarity (age, gender, major, native language)</td>
<td>Bodily Synchronization; Task Performance; Collaboration processes</td>
<td>Covary</td>
<td>Miles et al. (2011)</td>
</tr>
<tr>
<td>Social value orientation</td>
<td>Bodily Synchronization</td>
<td>Covary</td>
<td>Lumsden et al. (2012)</td>
</tr>
<tr>
<td>Lack of task knowledge</td>
<td>Task Performance; Collaboration processes</td>
<td>Covary</td>
<td>Fiorella, Vogel-Walcutt, &amp; Fiore (2012)</td>
</tr>
<tr>
<td>Cognitive load</td>
<td>Task Performance; Collaboration processes</td>
<td>Covary</td>
<td>Zheng &amp; Cook (2012)</td>
</tr>
<tr>
<td>Lack of communication between participants during the task</td>
<td>Task Performance; Collaboration processes</td>
<td>Included initial collaborative task to facilitate communicative exchange</td>
<td>Initial piloting of the experiment</td>
</tr>
<tr>
<td>Prior experience with Moonbase Alpha simulation</td>
<td>Task Performance; Collaboration processes</td>
<td>Control</td>
<td>N/A</td>
</tr>
</tbody>
</table>
CHAPTER THREE: METHOD

Design

At a general level, the aim of this research is to better understand verbal and non-verbal interaction dynamics of dyadic teams engaged in a CPS task and the capacity to use these measures to predict CPS performance. In order to ensure a sufficiently complex CPS task NASA’s Moobase Alpha simulation (NASA, 2011) was selected, given the necessity for collaborative interaction to achieve effective task performance. Research on interaction dynamics during dyadic team performance tends to utilize paradigms in which the natural variation of interaction is allowed to occur as opposed to attempting to artificially manipulate the sequence of the interaction, often through the use of confederates (e.g. Kenny, 1996; Louwerse et al., 2012).

Therefore, to meet the general aim of this research, a single interaction design was employed (cf. Malloy & Albright, 2001). In other words, this means that participants were paired up and interacted with each other in this experiment as dyads for a single primary interaction and did not interact with any other dyads. This design was selected because it allows for the natural emergence of interaction dynamics within the constraints of the CPS task. What becomes of interest in such designs, at least for the present purpose, is ultimately how the dynamics vary across dyads and the relationship those dynamics have with effective problem solving.
Participants

A total of 112 undergraduate participants were recruited to participate in this experiment comprising 56 dyadic teams. In order to participate, participants must have had general video game experience using a mouse and keyboard for third-person video games, no prior history of seizures, and no experience using the Moonbase Alpha simulation. Due to violation of these selection criteria or contamination of video files, a total of 13 teams’ data was excluded from further examination equaling a total of 43 dyadic teams (86 total participants) included in the present research (55 male, 31 female, $M_{age}$ = 19.2 years, range 18-28 years). Recruitment took place through the University of Central Florida’s Psychology department SONA system for which students who participated received 1.5 credits towards course requirements. Performance in the experiment was incentivized by offering a reward of $25 to each participant in the top-performing team.

Power Analyses and Sample Size

In this research, a combination of linear and nonlinear analytic techniques was used. Nonlinear methods are robust to violations of the assumptions of linear methods and specifically, recurrence quantification (described in the Analyses section) requires no assumptions about the data (Shockley, 2005). Power analysis does not apply in the traditional sense when combining nonlinear and linear analytic techniques. Therefore, a sample of articles employing similar methods was reviewed (i.e., those detailed in the Literature Review; see also Delaherche et al., 2012) and a conservative estimate of 40, on the high-end range of dyadic sample size (range: 1-50 dyads), was chosen to ensure there would be enough power to detect the desired effects. Thus,
the total inclusion of 43 dyadic teams in the analyses is justified by prior work. Tests of statistical power are also discussed in the Results section as appropriate.

Materials

Two desktop computers were set up so that participants sat face-to-face with each other with the computers offset slightly to one side (see Figure 2). Extant research has examined whether or not participants tend to synchronize when they have visual information about each other, with a majority of findings suggesting the phenomenon does occur with or without visual information (e.g., Shockley et al., 2003), but sometimes to a lesser degree without (e.g., Richardson et al., 2006). Further, pilot testing demonstrated that there is sufficient variation in dyadic movement for the proposed research. Another rationale for face-to-face interaction, was that it allows for participants’ computer screens to be placed back to back such that participants cannot view each other’s screens and thus strengthens the need for communication and coordination.

A Logitech HD webcam model C615 was used to record bodily movement of participants from a profile view (see Figure 2). All videos were collected in 720p HD resolution. 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. XSplit screen capturing software was used to record a video of participant performance in the simulation. Lastly, Qualtrics, an online survey system, was used to collect participant demographics and questionnaire responses.
In addition to the measures and apparatus described below, a set of paper-based materials were required to complete the Lunar Survival task (see APPENDIX A: LUNAR SURVIVAL MATERIALS for materials used on this task). The necessity of this initial task was determined as a function of the relative lack of communication between participants during piloting of the study. Therefore, this task served as a mechanism to increase non-acquainted participants’ conversational familiarity with each other and is not of further interest with regard to the research questions at hand. In short, the materials required for this task included a description of the scenario and description of the task, a list of the items and brief descriptions that participants need to rank in order of their utility for lunar survival, a form for participants to fill in their individual and collective rankings, and a form that shows the experts’ rankings of the items.

**NASA Moonbase Alpha Simulation Task Description and Complexity**

In order to engage the teams in a collaborative problem-solving task, NASA’s, mouse and keyboard-controlled, Moonbase Alpha simulation was used (NASA, 2011). This simulation places team members in a scenario where a meteorite 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. To the best of my knowledge, only one research experiment has been conducted using this simulation, which focused on the computerized analysis of agent performance during the task (Huntsberger, 2011). Therefore, the purpose of this section is to not only provide an overview of the task, but to also justify it as a sufficiently complex task that necessitates collaborative problem solving.

The general goal of the Moonbase Alpha task is for participants to fix and/or replace damaged components of the life support systems to restore oxygen to the settlement in 25 minutes or less. The major components that require repair include solar panels, power cables, couplers, a power distributor, and the life support system itself. Figure 3 represents the critical components that require repair and shows how the energy flows from one component to the next. 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 time-frame; however, some strategies are better than others.

![Figure 3. Schematic representing the Moonbase Alpha settlement and flow of energy.](image)
As previously discussed, complex problem solving tasks are dynamic, time-dependent, and characterized by complexity as a function of highly interconnected components without clear linear relationships (Quesada et al., 2005). The Moonbase Alpha task is dynamic in the sense that there are features of the task, such as the steady decay of oxygen or heat emissions necessitating use of a robot in certain areas, which change independently of the problem solver’s actions. The task is highly time-dependent in the sense that initial actions made early on can have a major effect on actions taken later in the task. For example, a team may choose to bring a portable toolbox to the area of the simulation in which they are working. This action will result in the later action of collecting the tool from the toolbox as opposed to traversing a greater distance back to the equipment shed. In turn, this could lead to a reduction in time taken to complete the task.

More generally, the task is complex because there are many required acts and informational cues that form the basis for how to restore the life support system of the moonbase (cf. Wood, 1986). To successfully restore the production of energy and oxygen to the moonbase, the affected solar panels, power cables and couplers, power distributor, and life support system must be repaired or replaced. Warnings and damage levels are displayed on the screen to inform the participants which settlement components have been damaged and require attention. The participants then must make decisions on how to properly act based on the types of cues they receive. Various tools, robots, and a moon rover are available for use to complete the task in the given amount of time. Selection of the appropriate tools to use on a given system component, and a variety of ways and sequences from which to replace or repair components is critical to completion of the task in the allotted time. In addition to being complex, the task requires
problem solving because a) participants will not have had prior exposure to this task and b) there is not a readily apparent solution for how to accomplish the goal of oxygen restoration. In other words, there is a suite of necessary acts that participants who are facing this problem for the first time will have to complete, but what each person does (or does not do), and when, is ultimately up to them to orchestrate (e.g., through information exchange and knowledge building).

For this reason and others, the problem solving task requires the collaborative efforts of both team members. That is, given the complexity of the task and the required subtasks (and the associated acts for each), there is a very low probability of completing the mission in the 25 minute 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. More nuanced specifics of the Moonbase Alpha simulation are detailed in APPENDIX B: MOONBASE ALPHA TRAINING POWERPOINT, which is the PowerPoint presentation used to train participants for the task.
Measures

Biographical Data

A basic, 19-item biographical data questionnaire was used to solicit participants’ age, gender, race, visual acuity, native language, major, grade point average, general video game experience, check that participants have not used the Moonbase Alpha simulation previously, and determine whether or not participants knew each other and if so, for how long.

Similarity Measure

A simple measure of similarity was computed based on the biographical measures of age, gender, race, major, and native language. The rationale here was to compute a team-level metric that could be assessed for covariation with performance (see Data Screening section). For each team, the aforementioned biographical measures were compared for each team member. If the team members had an exact match on a given biographical measure, then a value of 1 was added to the team’s similarity score. Thus, the values of this measure took on a theoretical range of 0-5, although the maximum observed value was 4. The higher the value of the similarity measure the more similar the team was in terms of their age, gender, race, major, and native language; the lower the score, the more dissimilar the team.

Performance Measures

Problem solving performance using the NASA Moonbase Alpha simulation was determined as a function of (a) the total time taken to restore life support and (b) the total percentage of oxygen restored. 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 and in which amount of time, but also those that did not complete the task but still performed to some degree in completing the task. In other words, the proposed performance measure allows quantification of teams’ problem solving outcomes, even when they do not complete the task.

This single performance variable ranging from 0-100 was computed by creating a scaling measure that ranks performance based on a combination of the time taken to complete the task and the percentage of oxygen restoration. The full amount of points (100) was awarded based on the minimum observed time taken to complete the task (e.g., 15 minutes) with 100% oxygen restoration. In contrast, the lowest amount of points (0) was awarded in cases where 25 minutes were spent on the task with 0% oxygen restoration. This was done by first rescaling the total time to complete the task in seconds into a 0-100 scale with the following formula:

$$100 \times \frac{\text{MaxObservedValue}}{\text{ObservedValue}} / \text{Range},$$

where 1500 was the max observed value and the range was 488. Importantly, this rescaling creates higher values for lower times. Next, these values were added to the total percentage of oxygen restored and then divided by two to put the values back on a 0-100 scale.

**Interaction Measures**

The novel contribution of this research is quantification of the interaction dynamics that occur during collaborative problem solving. Because of the dynamic nature of such interactions, these measures actually serve dual functions. These measures were assessed in order to (a) understand the types of interaction dynamics that occur during CPS and (b) see the degree to which characteristics of the interaction dynamics are predictive of problem solving performance. In this way, these variables serve multiple functions, depending on the particular analytic
technique that was employed (e.g., detecting whether synchronization actually occurred versus whether synchronization predicted performance). Two measures of interaction dynamics were assessed that correspond to non-verbal and verbal interaction, respectively: bodily movement time-series and team communications time-series. Each of these is discussed below.

**Bodily Movement Time-Series**

A frame-differencing technique (described in the Analysis section below) was used to extract a time-series representing the gross level of bodily movement for each participant. For the current task, this measure of bodily movement captures behaviors such as postural sway, gestures, adjustment of position, hand movements controlling the mouse, and shifting of the legs or feet. Because participants were required to use a mouse and keyboard for the task, the frequency of gestures was relatively low, but some deictic (i.e., gesture referring to a location or direction), iconic (i.e., gesture in shape of some item or route being discussed), and metaphorical (i.e., gestures that help explain a concept; Louwerse et al., 2012) gestures were observed. In general, this technique provides an objective measure of the amount of movement a given participant is exhibiting moment-by-moment over the duration of the task. To visualize these data, Figure 4 shows two bodily movement time-series derived from this study.

For bodily movement time-series, the values range from 0 to 1.0 which correspond to the degree of pixel changes between frames. Using the frame-differencing technique, Paxton and Dale (2013b) demonstrated that low values indicated little to no bodily movement and high values indicated more bodily movement. The movement scale is represented by the y-axis in Figure 4. These graphs are on a time scale indexed by the frame number. The video that these movement time-series were extracted from was approximately 25 minutes and sampled at a rate
of 8 Hz. This leads to the frame range from 0 – 12,000 shown on the x-axis in Figure 4. Therefore, each frame corresponds to 1/8 of a second. This figure illustrates, at a gross level, the amount of bodily movement occurring for each participant over the entire duration of the 25-min collaborative problem solving task. Note, however, that if teams completed the task in less than 25 minutes then the frame range will be proportionately shorter.

![Bodily Movement Time Series](image)

*Figure 4. Two bodily movement time-series for a dyadic team.*

**Communications Time-Series**

Team communications were transcribed and coded to represent team knowledge building processes as detailed below in the section on Data Coding. The 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 5) and is listed sequentially as it occurred during the interaction (e.g., Gorman, Cooke, Amazeen, & Fouse, 2012). To create a time-series from categorical codes, each code is assigned a value such as those in Table 5. In addition to just reflecting the categorical content assigned to a given utterance, the coding values can be doubled so as to also represent who was communicating. For the present purposes, the coding scheme has a range of values spanning 1-15. When accounting for which speaker, for example, coded values in the range of 1-15 correspond to team member one and values from 16-30 correspond to team member two.

An example of three discrete sequential time-series representing different aspects of an observed team’s verbal interaction is shown in Figure 5. Given that the literature is not clear with regard to the most informative type of communications time-series, the present research investigated all three. A majority of research examining team communication dynamics has focused on something like the first time-series shown in Figure 5, which is indicative of only the speaker (cf. Gorman et al., 2012a). The y-axis of the first graph represents, discretely (values of 1 or 2), who is speaking (either person 1 or person 2) across the duration of the task, which is represented by utterance number on the x-axis. This form of time-series will be referred to as the speaker time-series. The other two time-series are similar, with the only difference being what is represented on the y-axis. The second time-series represents only the sequential team knowledge building code uttered by the team, irrespective of the speaker drawing from the values shown in Table 5 for those on the y-axis; this one is labeled code time-series. The third time-series accounts for both the team knowledge building code and the speaker, with the values shown on the y-axis representing the team knowledge building values shown in Table 5. These values are
modified by have initial values of the code (1-15) representing one participant and, additively, by doubling the code values, the second participant for all those values above 15. Thus, this is the code*speaker time-series.

Because differential meaning about the interaction dynamics can be attributed to results stemming from any given one of these time-series (e.g., the coordination of speakers versus the coordination of knowledge building processes) and variability in the sensitivity as a function of an increasing range of values in the time-series (cf. Russell et al., 2012), each was included in the present research and their specific uses are detailed in the Results section. It should also be noted that these are discrete sequential time-series and although it is common to represent them with continuous lines, they are not, although it helps to visualize the interaction patterns.
In addition, testing Hypothesis 3 requires the generation of additional time-series for each team. The rationale for this is discussed in more detail in the Results section. The point here is to merely illustrate the form that these additional time-series take. What is apparent about the time-series discussed so far is that they represent the interaction of the team as a whole. Complementary interaction dynamics have been argued to be prevalent at this level of analysis and evidenced by comparing with similar metrics derived from time-series that represent solely
the individual team member’s communications (e.g., Fusaroli et al., 2014; Fusaroli & Tylén, 2015).

Therefore, four additional time series were generated for each team to represent the individual communicative contributions of each member. One the one hand, two of the time-series were derived from the *speaker* time-series and each represented whether the given team member was speaking or not. In other words, for each team a dedicated time-series was created for both members that indicates whether that member was speaking or not for a given utterance. For clarity, I refer to these as the *speaker 1* and *speaker 2* time-series. On the other hand, the other two time-series were derived from the semantic *code* time-series. Each of the two created time-series represent the specific semantic code for a given team member’s utterance provided they were speaking for said utterance. For clarity, I refer to these as the *code speaker 1* and *code speaker 2* time-series. In order to retain the temporal patterning and avoid artificial inflation of recurrence (see Recurrence Quantification Analysis section), null codes for each of these time-series (e.g., the absence of speaking) were given a value that was distinct from the other (e.g., Fusaroli & Tylén, 2015). For example, the *speaker 1* time-series received a value of 0 to indicate “not speaking” and the *speaker 2* time-series received a value of 3 to indicate “not speaking”. Thus, because the two time-series have different values for “not speaking” the null codes do not count as a recurrent point. However, both time-series received a value of 1 to indicate “speaking”, because this is the activity of interest for the recurrence analysis. Given that the *code*speaker time-series already reflects both speaker and semantic code, it was not possible to generate separate time-series for each individual team member.
In sum, each team’s communications were represented by a total of seven time-series. Three represented the overall interaction in a single time-series by indicating speaker, semantic code, or code and speaker. Four represented the individual communicative contributions of each team member with one for each team member indicating whether they were speaking or not, and if they were speaking, the semantic content of that utterance.

Subjective Measures

The primary measures of interest for the present purposes are those that quantify the interaction dynamics and problem solving performance. However, the literature suggests that certain individual differences may predict, for example, the emergence of interpersonal synchronization or other forms of coordination. The present research therefore examined individual differences in social value orientation and cognitive load. Each of these subjective measures is described in the following sections.

Social Value Orientation (SVO)

SVO is a measure used to differentiate between individuals who are predisposed to be pro-social, individualistic, or competitive. A categorization is made as a function of participants selecting six responses that align with a specific orientation. Given that Lumsden et al. (2012) found that individuals with a pro-social SVO tended to synchronize their bodily movements more so than those with a pro-self SVO, a 9-item measure of SVO was included in this research (Van Lange, 1999). Each item of this measure represents an independent trial of an economic exchange game in which the participant must decide how to allocate points amongst their self and some hypothetical other. For each item, the participant is required to select among three
options as to how they prefer to allocate the points. As an example, the participant may choose option A), in which he or she gets 500 points and the other gets 100 points, B), in which both get 500 points, and C), where the participant gets 550 points and the other gets 300 points.

Following the work of Lumsden et al. (2012), a categorization of individualistic or competitive was categorized as pro-self in the present work. Thus, data from this measure indicate whether someone has a SVO that is pro-social or pro-self (some participants were unclassifiable due to not having six or more responses allocated to a given category). See APPENDIX C: SOCIAL VALUE ORIENTATION ASSESSMENT for the items included in this measure. Importantly, this measure has shown to be a valid predictor of social behavior (Bogaert, Boone, & Declerk, 2008).

Cognitive Load Questionnaire

Given that cognitive load (i.e., the amount of mental effort elicited by a task) is an important factor that influences performance during complex problem solving (e.g., Zheng & Cook, 2012), a 3-item Cognitive Load Questionnaire (adapted from Paas, 1992; Zheng & Cook, 2012) was used. These items required participants to indicate the amount of cognitive effort required by the task, the degree of difficulty associated with the task, and the amount of frustration experienced during execution of the task on a 9-point Likert scale. The measure of cognitive load for each participant was computed by taking the average response across each of these three items with values taking on a possible range of 1 - 9. The purpose of including this measure was to account for individual differences in problem solving performance that can be attributed to variation in cognitive load responses. The items for this measure can be found in APPENDIX D: COGNITIVE LOAD RATING SCALE.
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 correctly answered 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. The rationale for including this assessment was two-fold. On the one hand, the assessment was used as a technique to motivate participants to carefully review the training materials. Additionally, the results of the assessment were used to account for individual variation in problem solving performance. The items included in the knowledge test are included in APPENDIX E: MOONBASE ALPHA KNOWLEDGE ASSESSMENT.

Procedure

Upon arrival to the laboratory environment, 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 since they would be working together as a team to complete the problem solving task. After this, participants were given an informed consent document to review. Participants were then asked to complete the biographical questionnaire and the 9-item measure of SVO.

Next, participants completed the Lunar Survival Task (e.g., Bottger & Yetton, 1987; Hall, Mouton, & Blake, 1963) as a means of prompting participants to converse on a task. In short, the
Lunar Survival Task provided participants with the hypothetical scenario that they were stranded on the moon and awaiting pick-up from a space-craft. From a list of items, participants prioritized those that are most relevant to surviving on the moon. Participants did this by first rating the items individually and with explicit direction to think of a rationale for why some items are more important than others. Then, the team was tasked with developing a collective prioritization list. That is, they had to communicate and reason with each other so as to come up with a final list of ranked items. Participants were then provided with the expert ratings so they could see how their performance compared to the ideal responses (see APPENDIX A: LUNAR SURVIVAL MATERIALS for materials used for Lunar Survival Task). This task formed a foundation from which team members have some familiarity conversing with each other in a problem solving context to facilitate later communication during the Moonbase task.

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 and can be viewed in APPENDIX B: MOONBASE ALPHA TRAINING POWERPOINT. 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 experimenter reiterated the necessity for team members to communicate with regard to the required actions needed to complete the task within the allotted time. Participants were then prompted to put on the audio headsets and directed on how 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 minute task performance trial. However, following completion of the introduction video, the lights in the room were flashed to allow for temporal synching across video files. Participants were told to begin the task as soon as the lights turned back on. The Moonbase task was considered completed either when time ran out or once participants fully restored oxygen to the moonbase, whichever came first. Following completion of the task, participants completed the Cognitive Load Questionnaire. 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.

Communications Data Coding

All dialog was time-stamped for onset of a given utterance, speaker demarcated, transcribed, unitized, and coded using a coding scheme based on the MiTs framework (for similar methods see Fiore et al., 2014; Rosen, 2010; Srnka & Koeszegi, 2007). The coding scheme, originally developed by Rosen (2010) to capture team knowledge building processes, was used as the initial coding scheme from which 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 5. A detailed codebook was developed specifically for this task and includes more detailed definitions as well as conditions for when and when not to use a code. Examples can be found in APPENDIX G: COLLABORATIVE PROBLEM SOLVING CODEBOOK.

Reliability was established for the unitization process using Guetzkow’s U (Guetzkow, 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 (falling primarily under the team knowledge building component), Cohen’s Kappa (\( \kappa \)) 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), \( \kappa \) 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 (see APPENDIX G: COLLABORATIVE PROBLEM SOLVING 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 (κ = .67) and by the completion of the training reliability was excellent (κ = .81; Banerjee, Capozzili, 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 similarly 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 amongst 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).
Table 5
Team knowledge building codes to be used for communications data (adapted from Rosen, 2010).

<table>
<thead>
<tr>
<th>Process</th>
<th>Code</th>
<th>Brief Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Team Information Exchange</td>
<td>1. Information Provision (IP)</td>
<td>- 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.</td>
</tr>
<tr>
<td></td>
<td>2. Information Request (IR)</td>
<td>- Question utterances asking for a response of simple information about the task environment or situation, or questions asking for repetition of immediately preceding information.</td>
</tr>
<tr>
<td></td>
<td>3. Knowledge Provision (KP)</td>
<td>- Statements about the task environment or situation that provide either 1) an integration of more than one piece of simple information, or 2) an evaluation or interpretation of the meaning, value, or significance of information with regard to the current subtask.</td>
</tr>
<tr>
<td>Team Knowledge Sharing</td>
<td>4. Knowledge Request (KR)</td>
<td>- Question utterances that request a complex information response about the task environment or situation: to answer the question, the response should provide either 1) an integration of more than one piece of simple information, or 2) an evaluation or interpretation of the meaning, value, or significance of information within the current subtask.</td>
</tr>
<tr>
<td>Team Solution Option Generation</td>
<td>5. Option Generation – Part (OG-P)</td>
<td>- 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.</td>
</tr>
<tr>
<td></td>
<td>6. Option Generation – Full (OG-F)</td>
<td>- 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.</td>
</tr>
<tr>
<td>Team Evaluation and Negotiation of Alternatives</td>
<td>7. Solution Evaluation (SEval)</td>
<td>- Utterances that 1) compare different potential solutions, 2) provide support, criticism, or indifference to a potential solution, or 3) ask for evaluation of a solution.</td>
</tr>
<tr>
<td></td>
<td>8. Goal/Task Orientation (GTO)</td>
<td>- 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.</td>
</tr>
<tr>
<td>Team Process and Plan Regulation</td>
<td>9. Situation Update (SU)</td>
<td>- Statements that provide information regarding what the team is currently doing or what is currently happening with the simulation.</td>
</tr>
<tr>
<td></td>
<td>10. Situation Request (SR)</td>
<td>- Statements that ask about what the team is currently doing or what is currently happening with the simulation.</td>
</tr>
<tr>
<td></td>
<td>11. Reflection (R)</td>
<td>- Utterances that provide or ask for a critique or evaluation of the performance of the team as a whole or of individual members.</td>
</tr>
<tr>
<td></td>
<td>12. Simple Agree/Disagree/ Acknowledge (S)</td>
<td>- Simple agreement/disagreement utterances are expressions of agreement or disagreement with no rationale provided. Acknowledgements are utterances providing recognition of receipt of communication.</td>
</tr>
<tr>
<td></td>
<td>13. Incomplete/ Filler/ Exclamation (INC/F/EX)</td>
<td>- 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.</td>
</tr>
<tr>
<td></td>
<td>14. Tangent/Off-task (T/OT)</td>
<td>- 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.</td>
</tr>
<tr>
<td></td>
<td>15. Uncertainty (UNC)</td>
<td>- Uncertainty statements explicitly express either general or specific uncertainty about the roles, tasks, situations, or anything else task-related.</td>
</tr>
</tbody>
</table>
Data Analyses

Traditional measures of central tendency tend to be inadequate for capturing social and team interaction dynamics (e.g., Gorman et al., 2010). That is, because patterns of interaction change and fluctuate, a mean and standard deviation for a given behavioral or cognitive metric aggregated across some period of time, fails to capture the dynamics that occur (i.e., they are often non-stationary; Dale, Warlaumont, & Richardson, 2011). To address this issue, researchers interested in studying interaction have increasingly relied on the use of time-series analysis; often using nonlinear analytic techniques. In general, this means a given variable of interest is sampled at a regular time interval across the entirety of the task duration. In the present research, both continuous and discrete time-series analyses are employed.

Continuous time-series analyses were derived by using a frame-differencing technique to extract the degree of bodily movement exhibited by each participant, which is described in the next sub-section. In contrast to a continuous time-series, which is sampled at a regular time interval, a discrete time-series is often used to characterize a set of nominal, mutually exclusive codes in the sequence that they occur, without having a particular time interval associated with it (Gorman et al., 2012a). The discrete sequential time-series employed in the present research is composed of the collaborative communications coding scheme detailed above and will be subject to recurrence quantification analysis, which is detailed in a sub-section below.

Frame-Differencing Technique

To analyze interpersonal synchronization, Paxton and Dale’s (2013b) frame-differencing method was utilized. This technique captures the degree of gross level bodily movement
exhibited by interacting pairs. First, video data collected from the experiment was segmented into frames at an 8 Hz sample rate using the Free Video to JPEG Converter. The number of desired frames was manually specified for input into this program. Due to variation in total video length as a function of start time, the total number of frames was determined by converting the video length into seconds and then multiplying this number by eight in order to satisfy the 8 Hz sampling rate (i.e., 8 frames were collected for each second of video). More formally, \( N(\text{#Frames}) = T(\text{Time length of video in seconds}) \times 8(\text{Frames/second}) \).

A maximum of approximately 12,000 frames was captured for each dyad after truncating the frames collected prior to or after participants began performing the moonbase task. The frame-differencing method, applied using the Matlab script provided by Paxton and Dale (2013b), sequentially loads each frame and halves it to separate the two participants comprising the dyad. Over the course of loading and processing each frame, the method compares pixels of the respective half of the frame with the prior half frame to compute a raw pixel change score, which is then transformed into a standard difference score. Ultimately, two time-series are derived in this way such that they represent the respective participant on the right or left half of the frame. Assuming all else is held static, the values generated here correspond to the gross amount of bodily movement for each participant over the duration of the task. A second-order Butterworth low-pass filter was applied to each of these movement time-series to control for any lighting fluctuations and other noise while remaining sensitive to slight bodily-movement fluctuations (Paxton & Dale, 2013b).

Next, the two standardized time-series are subject to a time lagged cross-correlation analysis, which is often used as a measure of behavioral coordination (Richardson et al., 2014).
That is, the two time-series are cross-correlated first with a time lag of 0, and a value for Pearson’s correlation coefficient $r$ is obtained for the time-series at that time. Then, depending on the specified lag time, the analysis would compute the cross correlation at lag of -1, which means that the left participant’s time-series at time $t$ is compared to the right participant’s time-series shifted a step to $t + 1$. Conversely, a lag of +1 cross-correlates the time-series of the left participant’s movement at $t + 1$ with the time-series of the right participant at time $t$. A figure is helpful in explaining this. Figure 6 shows an example from the data representative of this analysis. Prior research has shown that when the peak cross-correlation is at a lag time of 0, then this suggests that the two participants’ movements are indeed synchronized (Paxton & Dale, 2013a, 2013b). If the peak cross-correlation is to the left or right of time lag 0 then this is indicative of a leader-follower type relationship (see also Boker, Rotondo, Xu, & King, 2002).

![Bodily Movement Lagged Cross-Correlation Plot](image)

*Figure 6. Example time lagged cross-correlation analysis plot.*

To determine that the frame-differencing method is indeed a valid measure of bodily synchrony, Paxton and Dale (2013b) conducted a comparison test assessing the output of the frame-differencing method with that of a human rater, a more traditional means of assessing synchrony (Bernieri & Rosenthal, 1991). Specifically, a scripted interaction meant to display
bodily synchronization was recorded. A human rater then went through and observed the movement for each individual in the video independently. For each one-second interval, the rater indicated, using a 1-7 point Likert scale, the overall bodily movement present in that second. The cross-correlation function from this analysis was then compared to that derived using the frame-differencing technique. Results showed that the two methods were strongly correlated with $r = .68, p < .0001$. Further, a similar comparison technique was done using a subset of the actual data presented in Paxton and Dale (2013b) with congruent findings supporting the validity of the frame-differencing technique as a measurement of bodily synchrony and not some other artifact captured in the video. The aim here was to apply this video frame-differencing method to examine interpersonal synchronization during collaborative problem solving. However, given that time-lagged cross-correlation analysis is a linear technique, a test of stationarity was warranted to further determine its appropriateness (see Data Screening section).

Recurrence Quantification Analysis

*Recurrence*, defined as the degree to which a system comes to inhabit the same or a very similar state, is a fundamental property of dynamical systems (Marwan, Romano, Thiel, & Kurths, 2007; Webber, & Zbilut, 2005). Originally applied to the study of physical systems, recurrence-based analyses have increasingly been applied to the study of the dynamics of social systems (e.g., Dale & Spivey, 2006; Richardson et al., 2014; Shockley, 2005) and team interactions (e.g., Gorman et al., 2012; Russell et al., 2012). The guiding motivation behind this analysis is that it allows for the quantification of a system’s dynamics by ascertaining whether behavioral states of the system tend to recur over time and, if so, the patterns in which these recurrences are structured, and how complex and flexible they are (e.g., Fusaroli, Konvalinka, &
Wallot, 2014; Richardson et al., 2014). Importantly, it has been argued that recurrence quantification provides an objective measure of the temporal dynamics of cognitive and behavioral processes depending on the phenomenon under study (Richardson, Dale, Shockley, 2008). This technique is important for the present purposes as human interactions tend to unfold in non-stationary and temporarily non-independent ways, and it is robust to the detection of various forms of multiscale coordination in all manner of interpersonal behavioral modalities (see Fusaroli et al., 2014; Louwserse et al., 2012).

The recurrence plot (RP) developed by Eckmann, Kamphorst, and Ruelle (1987) was originally established as a technique to qualitatively examine the dynamics of a system (see Figure 7 for example recurrence plots). Put simply, a recurrence plot represents the set of values in a time-series with numeric identifiers that have a distance of 0 from each other and indicates when a system revisits the same state over time (Marwan et al., 2007). More formally, “Recurrence plots are symmetrical $N \times N$ arrays in which a point is placed at $(i, j)$ whenever a point $X_i$ on the trajectory is close to another point $X_j$” (Zbilut & Webber, 2006, p.1). In an over simplification that helps to visualize this plot, a recurrence plot can be thought of as pairing a time-series with itself, so if a given state occurs at time 3 and it recurs at time 7, then a recurrent point will be placed on the plot at the intersection of where time 3 occurs on one axis and time 7 occurs on the other.

Generating a recurrence plot relies on phase space reconstruction. Phase space is ultimately characterized by the variables that define the state that a system is in, with each variable being one dimension of that space. Further, each point in phase space corresponds to a certain state of the system. In social systems, as with most systems, only so many variables of a
system can be observed. Phase space reconstruction, then, is a technique based largely on Takens’ (1981) theorem that allows information about the entire system’s behavior to be gleaned from only a single observed variable (i.e., one dimension). Takens (1981) proved that if a system is nonlinear, the topology of how that system moves through phase space can be reconstructed using time-lagged, yet exact, copies of the time-series of the observed variable.

The precise formalisms and proofs supporting this technique are beyond the scope of the present work and can be found elsewhere (e.g., Webber & Zbilut, 2005). The point here is to acknowledge the underlying mechanics of generating a recurrence plot. In particular, current techniques for phase space reconstruction rely primarily on the ability of the researcher to appropriately set three parameters: embedding dimension, delay, and radius. Determining the appropriate parameters differs as a function of whether the time series is categorical or continuous. Because parameter identification is easier to understand for categorical data, such examples are provided in the following paragraphs. However, phase space reconstruction involves additional tests when applied to continuous data, which are beyond the present scope (see, however, Aks, 2011; Webber & Zbilut, 2005 for thorough treatments).

The embedding dimension (M), in the case of categorical data, can be thought of as a parameter that defines how many sequential states must match for the recurrence plot to generate a recurrent point (Dale, 2014; Zbilut & Webber, 2006). So, in the case of categorical data it is typical to select an embedding dimension of $M = 1$ so that a category match counts as a recurrent point each time said category recurs in the time-series (Dale et al., 2011), as opposed to requiring a certain sequence of categories. For example, using the current coding scheme (see Table 5), each time an instance of team information provision occurs, after having already occurred once,
a point will be placed on the recurrent plot. When \( M = 2 \), for example, then if team information exchange and team knowledge sharing occurred sequentially after having already occurred together, a point would be placed on the recurrent plot.

The *delay* \( (\tau) \) specifies the time lag at which the time-series, in whatever temporal units that time-series is characterized by, will be shifted for comparison with itself (i.e., auto-recurrence; Webber & Zbilut, 2005). With categorical data, it is best to preserve the temporal order (Dale, 2014) and thus \( \tau = 1 \) is the optimal selection for this parameter as there are “no points in the time-series that are skipped” (Webber & Zbilut, 2005, p. 37). That is, as \( \tau \) increases, points in the time-series are truncated.

The *radius* is a parameter that specifies the distance from a given value in the time-series that will be counted as a recurrent point (Dale, 2014). For example with a radius of 1, any values in the time-series that have a numerical distance of one will count as a recurrent point. For example, in the coding scheme shown in Table 5, the values for information request and knowledge provision (values 2 and 3, respectively) would interchangeably count as a recurrent point even though categorically they are very distinct. For this reason, in cases of categorical time-series, the radius is often set at or very near to zero so that only an exact category match will count as a recurrent point (Dale, 2014).

To make the idea of recurrence plots somewhat less esoteric, three different recurrence plots were generated using the same three time-series shown in Figure 7, which represent the *speaker*, *code*, and *code*\(^*\)speaker time-series. What is immediately noticeable are the differences in the density of recurrent points in these plots with the *speaker* time-series having the most and the *code*\(^*\)speaker time-series having the least. Considering that the range of possible states spans
2, 15, and 30 for each respective time-series, it is an intuitive interpretation that it is easier for a system to inhabit the same space more frequently (higher recurrence) when there are a fewer number of possible states. An important point to mention here is that recurrence plots are symmetrical along the diagonal, which is referred to as the line of identity (LOI) and is typically excluded from any recurrence quantification analyses (Marwan et al., 2011). Besides this main diagonal, any subsequent points that fall on a diagonal line represent sequences of recurring states. Indeed, examining the diagonal structures is a major aspect of recurrence quantification analysis (RQA) in addition to the amount and distribution of points.

![Figure 7](image)

**Figure 7.** Three example recurrence plots generated from the *speaker, code, and code*\(^*\)speaker time-series shown in Figure 5.

With the need to quantify and derive greater meaning from these plots, Webber and Zbilut (1994) developed recurrence quantification analysis (RQA). Given the symmetry inherent to RPs, RQA focuses on quantifying only the information in the upper triangle as it is redundant with the lower triangle (Webber & Zilbut, 2005). A variety of measures have been developed
that quantify the RP and provide some index of the dynamics of the system under study. Only the most relevant RQA measures to the present research are detailed here: recurrence rate (%REC), percent determinism (%DET), average diagonal line length (MEANLINE), and entropy (ENTROPY).

%REC quantifies the percentage of the plot that is occupied by recurrent points (Dale, 2014; Zbilut & Webber, 2006). Values of this measure can range from 0%, meaning there are no recurrent points evident in the plot to 100% in which all points are recurrent points (Webber & Zbilut, 2005). In the most general sense, this measure indicates how frequently a system revisits states that it was in previously. In the context of the current categorical coding schemes, depending on which of the time-series is used (see Figure 5), this measure would provide an index of how much problem solving processes recur by the team as a whole or by specific team members. More generally, it can be used as a metric to differentiate teams who are exhibiting many of the same behaviors throughout the collaborative problem solving task, from those who are exhibiting relatively diverse communicative behaviors.

%DET quantifies the percentage of points in the RP that fall on a diagonal line compared to the total number of recurrent points (Dale, 2014; Zbilut & Webber, 2006). Generally, this measure is used to differentiate whether or not the system exhibits periodic (many long diagonal lines and high %DET), chaotic (many short diagonal lines and low %DET), or stochastic behavior (few to no diagonal lines at all and very low to 0 %DET). In the team coordination context, high %DET has been argued to be indicative of a team that is performing rigidly; whereas, lower %DET has been argued to be indicative of a team that is performing flexibly (cf. Gorman et al., 2012a).
MEANLINE quantifies the average length of the diagonal structures evident on the plot (Dale, 2014). This is an important index because it indexes the average length of interaction sequences that a team may exhibit multiple times and thus can reflect the stability of the interaction.

ENTROPY quantifies the Shannon information entropy for all the diagonal line lengths across their distribution in a histogram (Webber & Zbilut, 2005). In general, this metric provides an index of the complexity of interaction sequences. That is, given diagonal lines in the RP represent recurrent sequences of certain interaction patterns, a team that has higher variety in the lengths of these interaction sequences would be characterized as more disorderly or chaotic (high ENTROPY); whereas, a team that was consistent in their length of recurrent interaction patterns would be characterized by more regularity in their interaction dynamics (low ENTROPY). Table 6 summarizes the measures provided here and the potential insight they can offer when applied to team communications during CPS.
Table 6

Summary of RQA measures.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Definition</th>
<th>Meaning in context of team communications during CPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent Recurrence (%REC)</td>
<td>Percentage of the plot that is occupied by recurrent points</td>
<td>A metric to differentiate teams exhibiting many of the same behaviors from those who are exhibiting relatively diverse behaviors</td>
</tr>
<tr>
<td>Percent Determinism (%DET)</td>
<td>Percentage of points in the RP that fall on a diagonal line compared to all other number of recurrent points</td>
<td>High %DET is indicative of a team that is performing rigidly; whereas, lower %DET is indicative of a team that is performing flexibly and adaptively</td>
</tr>
<tr>
<td>MEANLINE</td>
<td>Average length of the diagonal structures evident on the plot</td>
<td>Average length of interaction sequences that a team may exhibit multiple times</td>
</tr>
<tr>
<td>ENTROPY</td>
<td>Shannon information entropy for all the diagonal line lengths across the bins they are distributed amongst in a histogram</td>
<td>Quantifies the complexity of the interaction sequences of the team, with higher entropy values indicating more disorderly interactions and low entropy meaning more structured or regular interactions</td>
</tr>
</tbody>
</table>

Up until this point, all discussion of RQA has focused on understanding the dynamics of a single univariate system (i.e., auto-recurrence). This method easily scales to understanding the coupling between two systems using cross-recurrence quantification analysis (CRQA). As should be clear, analyzing the whole team’s communications with RQA can be very informative. However, quantifying the recurrent structure between each individual team member’s communications, can also provide some unique insights. Generally, CRQA can examine how often two systems tend to visit the same states, whether or not they exhibit similar behavioral patterns, and how synchronous those two systems are operating (Fusaroli et al., 2014).

Much of the recurrence plot-based measures work the same in CRQA as they do in RQA, but the meaning behind them is slightly different. For example, %REC now provides a measure
of how often the two systems revisit the same states. %DET provides the proportion of recurrent points that form diagonal line structures, which in CRQA indicates that both systems are inhabiting sequences or patterns of behavioral states either at the same time or at different times. MEANLINE then conveys the average length of sequences of recurrent behavioral states. Lastly, ENTROPY quantifies the complexity or regularity of the interaction between the two systems (Fusaroli et al., 2014).

With CRQA, there is an additional measure not present in RQA. Recall that in RQA, there is always a long diagonal line in the center of the plot because the time-series is plotted against itself (i.e., the line of identity). This line is always there because this center diagonal represents a delay of 0, so the two time-series are identical at that time point. In CRQA, however, the diagonal becomes the line of coincidence or synchronization because, if there is a long diagonal in the center of a cross recurrence plot, then this means the two systems are behaving more or less the same at the same time (e.g. Coco & Dale, 2014a). For this reason, in CRQA, the diagonal recurrence profile (DiagProfile) becomes of interest to quantify. Essentially, this technique quantifies the amount of recurrent points at a delay of 0 and +/- some window size. The output from this analysis is very similar to that shown in Figure 6 where if the peak of recurrent points is at a delay of 0 than the system is synchronized; whereas, if the peak is at some other delay, it is evidence of a leader-follower dynamic (e.g., Fusaroli, 2014).

In sum, RQA and CRQA, as well as the frame-differencing method, serve as a foundation for the analytic techniques that were employed in this research investigating the interaction dynamics inherent to collaborative problem solving teams. The analyses conducted to test the hypotheses of the present research as well as the results are detailed in the next chapter.
CHAPTER FOUR: RESULTS

A set of linear and nonlinear analytic techniques were employed for the present research. Statistical analyses were performed using a variety of software including Matlab 2014, R version 3.1.2, and SPSS version 21.0. The respective software used is specified for a given analysis. An alpha level of $\alpha = .05$ was used unless otherwise noted.

Data Screening

Normative techniques for screening data were conducted including checking for outliers, missing data, and checking for normality. As appropriate, any screening techniques that modified the data are detailed with the respective analysis.

Assessing Potentially Confounding Variables

An important aspect of data screening for the present research was evaluation of those variables that were deemed as potentially confounding to testing the hypotheses (see Table 4). Screening for whether these variables should be examined further was based on evaluation of the bivariate correlation matrices calculated and shown in APPENDIX F: CORRELATION MATRICES. Specifically, three correlation matrices are shown in APPENDIX F: CORRELATION MATRICES corresponding to individual, team, and cross-level variables. Variables that were non-significantly correlated with performance, or that were redundant were removed from further consideration as covariates (see Becker, 2005).

The major hypotheses of this study focus on examining the emergence of certain patterns of interaction of teams and the effects those team interaction dynamics have on problem solving performance. This is important because these variables reside at the team level (i.e., team
performance, interpersonal synchronization). However, in Table 4, many of the variables that could confound the results are at an individual level. One way for addressing variables at multiple levels is to use multi-level modeling (see Dedrick et al., 2009 for review). Because the focus of the present study is on the relationship between variables residing at the team level, testing for cross-level interaction effects between the covariates and team-level metrics using multi-level modeling was beyond the scope of the present study and further, it would also require a larger sample size to gain sufficient power (e.g., Mathieu, Agunis, Culpepper, & Chen, 2012).

Given that multi-level modeling was not appropriate for the present purposes, a valid method for creating a team-level variable from individual-level variables was adopted (Chen, Mathieu, & Bliese, 2004). Specifically, a selected score model approach was chosen with a particular focus on participants’ video game familiarity, knowledge assessment scores, and cognitive load. These were selected given their establishment as important factors in related research, their observed correlation with performance, and non-redundancy with other measures. The goal of a selected score model is to determine what individual level metric is most representative and/or influential at the aggregate (i.e., team) level. For each dyad, the average, minimum, and maximum scores were computed for video game familiarity, knowledge assessment scores, and cognitive load. To determine which of these had the strongest relationship with the team-level variables, bivariate correlations were conducted. The results of this analysis (shown in APPENDIX F: CORRELATION MATRICES) suggested that many of these variables were correlated significantly with performance, but not synchronization or complementarity (%REC of code time-series see Testing Hypothesis 3 section). Thus, the versions of each variable with the highest correlation values were selected in order to control for
the most variance that could be attributed to those variables in subsequent regression analyses. In particular, the minimum amount of video game familiarity (MinVidGame), the minimum score on the knowledge assessment (MinKnow), and the average cognitive load (AveCogLoad) held the highest correlations and were included as control variables when appropriate and detailed in the analyses that follow.

Assessing Stationarity of Bodily Movement Time-Series

Linear time-series based analytic techniques (e.g., time-lagged cross-correlation) assume stationarity of the data. This simply means that it is assumed that measures of central tendency are accurate summarizations of the time-series; however, if the data are nonstationary, then measures of central tendency will vary as a function of where in the time-series it is evaluated and with what window size (Webber & Zbilut, 2005). Therefore, the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test was conducted using R to screen each of the bodily movement time-series extracted using Paxton and Dale’s (2013b) frame-differencing technique. This test assesses the null hypothesis that the time-series in question is stationary ($\alpha = .01$; see Strang et al., 2014). The rationale for employing this test was to determine whether the bodily movement time-series can be subject to linear time-series analysis (i.e., time-lagged cross correlation) or if they should be subject to the nonlinear, recurrence quantification analysis, which does not assume stationarity. The KPSS test was run on all participants’ bodily movement time-series ($n = 86$). A total of 0 of the 86 participants had $p$-values less than .01 suggesting that a majority of the time-series were stationary. For this reason, the linear, time-lagged cross correlation analysis was deemed appropriate to conduct on the bodily movement time-series.
Testing Hypothesis 1: Emergence of Interpersonal Synchronization

After assessing the degree to which the bodily movement time-series were stationary, these data were subject to time-lagged cross correlation analysis to extract the degree of interpersonal synchronization. In particular, the R CCF function was used to calculate cross correlation values at +/- 150 lag points. This roughly corresponds to approximately 18 seconds in either direction. Points at which \( r \) values are higher indicates whether the movements are changing in the same direction at the same time (values close to lag 0), whether the left participant is leading the right participant (negative lags), or whether the right participant is leading the left participant (positive lags). In order to derive a metric suitable for testing the first hypothesis, the mean cross-correlation values for +/- 4 lag points was computed for each team. This corresponds to roughly a 1-second window in which a half second of lag is allowable in either direction. This is a conservative estimate that is examining only short-term dependence between two participants’ bodily movements (cf. Marmelat & Delignières, 2012). Observed values ranged from 0.019 - 0.30. Recall that synchronization was operationalized for the present research as participants’ moving their bodies at approximately the same time. Therefore, it follows that the mean cross-correlation value of +/- 4 lag points validly represents the degree to which participants were moving their bodies at the same time.

However, in order to verify that synchronization cannot be attributed to chance, surrogate analyses are required (Ramseyer & Tschacher, 2010). To determine that the synchronization observed in dyads was not due to chance alone, a baseline level of synchronization was computed by shuffling each of the observed bodily movement time-series for each participant one hundred times using the R surrogate function in the tseries package (Hornik, 2014). This
technique is based on the algorithm developed by Theiler, Galdrikian, Longtin, Eubank, and Farmer (1992) and it destroys the temporal structure of the data while keeping distributional properties of those data intact (e.g., Louwerse et al., 2012). Then, the mean cross-correlation value for +/- 4 lag points across each of the 100 surrogates was computed. These values thus provide a surrogate or “pseudosynchrony” value for each team that corresponds to the average amount of synchronization expected by chance across 100 pseudo-observations for each team (see Ramseyer & Tschacher, 2010).

From this, an independent-samples, one-tailed t-test \((N = 86)\) was conducted using SPSS to examine whether the observed synchronization was significantly greater than the baseline level that could be expected due to chance. This analysis serves as a test of Hypothesis 1. Results showed that observed synchronization \((M = .088, SD = .05)\) was significantly greater than the surrogate or baseline level of synchronization \((M = .0006, SD = .002)\), \(t(84) = 11.22, p < .0001\). Further, there was a large effect size, \(r = .77\).

Figure 8, shown below, demonstrates the difference between observed synchronization and the surrogate-baseline synchronization. The black time series was taken from the observed cross correlation values of one team. Note the peak of the values around the lag 0. The other four lines in the figure represent the cross-correlation values for several surrogate series. This example clearly illustrates a case where observed synchrony was greater than the surrogate synchrony. Not all observations were as clearly distinct from the surrogates. Nonetheless, the results of the t-test support Hypothesis 1, suggesting that interpersonal synchronization, in the bodily movements of participants, did emerge during dyadic collaborative problem solving beyond a baseline level of synchronization due to chance alone. Therefore, for all subsequent
analyses, the values for interpersonal synchronization were represented by the mean cross-correlation value for +/- 4 lag points for a given team.

*Figure 8.* Example comparison of observed cross-correlation values versus surrogate cross-correlation values for one team.

**Testing Hypothesis 2:** Does interpersonal synchronization predict team problem solving outcomes?

A hierarchical linear multiple regression model was used to examine the relationship between interpersonal synchronization and problem solving performance while controlling for video game familiarity, knowledge assessment scores, and cognitive load using SPSS version 21.
Based on the aforementioned correlations of individual and team-level control variables, the minimum value for a dyad’s video game familiarity (MinVidGame), the minimum score of the dyad on the knowledge assessment (MinKnow), and the average value of each dyad’s cognitive load (AveCogLoad), were included as the control variables and thus entered into Step 1 of the model. The predicted variable was problem solving performance on the Moonbase Alpha collaborative problem solving task (see Method for how this was computed). Taken together, the control variables MinVidGame, MinKnow, and AveCogLoad accounted for a significant 53% of the variance ($R^2 = .53$, $R^2_{adj} = .50$), $F(3, 39) = 14.89$, $p < .0005$. All control variables were uniquely significant predictors of performance: MinVidGame ($sr^2 = .16$, $t(39) = 3.68$, $p < .01$), MinKnow ($sr^2 = .10$, $t(39) = 2.92$, $p < .01$), AveCogLoad ($sr^2 = .16$, $t(39) = -3.61$, $p < .01$).

In Step 2, interpersonal synchronization was added as a predictor of performance. The addition of synchronization as a predictor added a non-significant 1.3% of the variance to the model, $\Delta R^2 = .013$, $F_{\Delta R^2}(1, 38) = 1.06$, $p = .31$. Taken together, interpersonal synchronization, MinVidGame, MinKnow, and AveCogLoad accounted for 55% of the variance in performance, ($R^2 = .55$, $R^2_{adj} = .50$), $F(4, 38) = 11.45$, $p < .0005$. In this model, again all control variables were uniquely significant predictors of performance: MinVidGame ($sr^2 = .17$, $t(38) = 3.79$, $p < .01$), MinKnow ($sr^2 = .08$, $t(38) = 2.66$, $p < .05$), AveCogLoad ($sr^2 = .16$, $t(38) = -3.62$, $p < .01$). However, synchronization was not a uniquely significant predictor of performance, ($sr^2 = .01$, $t(38) = 1.03$, $p = .31$).

Table 7 summarizes the results of this hierarchical multiple linear regression analysis including the standardized coefficients. Assumptions of the model were assessed in two ways.
First, visual inspection of the residual plots and histogram did not reveal any clear deviations from homoscedasticity or normality. Further, the \texttt{gvlma} function in R was used as a global test of the model’s assumptions. Results from this test showed that the assumptions of the model were not violated, thus suggesting it was an appropriate model of the data.

Thus, the current results do not show support for H2 in that interpersonal synchronization did not predict performance on the Moonbase Alpha collaborative problem solving task. A post-hoc power analysis was conducted to ensure the data were sensitive to detecting any effect synchronization may have had on performance. Specifically, \(1 - \beta = .69\), with the observed effects size \((R^2) = .11\) and the observed \(\alpha = .31\). Thus, it is reasonable to conclude that the odds were favorable that the analyses had enough power to detect any effect that synchronization may have had on performance.

Table 7
\textit{Synchronization predicting performance multiple regression output.}

<table>
<thead>
<tr>
<th>Step</th>
<th>Variable</th>
<th>B</th>
<th>SE B</th>
<th>(\beta)</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Constant</td>
<td>-48.44</td>
<td>35.77</td>
<td>-1.35</td>
<td>.18</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MinVidGame</td>
<td>14.01</td>
<td>3.80</td>
<td>.410</td>
<td>3.68</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>MinKnow</td>
<td>10.11</td>
<td>3.46</td>
<td>.323</td>
<td>2.92</td>
<td>.006</td>
</tr>
<tr>
<td></td>
<td>AveCogLoad</td>
<td>-9.65</td>
<td>2.67</td>
<td>-.399</td>
<td>-3.61</td>
<td>.001</td>
</tr>
<tr>
<td>2</td>
<td>Constant</td>
<td>-50.20</td>
<td>35.78</td>
<td>-1.40</td>
<td>.169</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MinVidGame</td>
<td>14.56</td>
<td>3.84</td>
<td>.426</td>
<td>3.79</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>MinKnow</td>
<td>9.40</td>
<td>3.53</td>
<td>.300</td>
<td>2.66</td>
<td>.011</td>
</tr>
<tr>
<td></td>
<td>AveCogLoad</td>
<td>-9.67</td>
<td>2.67</td>
<td>-.400</td>
<td>-3.62</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>Synchronization</td>
<td>71.45</td>
<td>69.78</td>
<td>.115</td>
<td>1.03</td>
<td>.310</td>
</tr>
</tbody>
</table>
Testing Hypothesis 3: Complementarity in team communicative interaction sequences

To address Hypothesis 3 and to provide the measures needed to test Hypothesis 4, a series of analyses were performed on the teams’ overall communications time-series (i.e., speaker, code, and code*speaker; see Figure 5) as well as the individual communicative contributions time-series (i.e., speaker 1, speaker 2, code speaker 1, and code speaker 2). The rationale for the analyses detailed here is that RQA at the level of the dyad as a whole, has been argued to be more sensitive to detection of complementary interaction dynamics and predictive of performance on joint collaborative tasks, particularly when compared to CRQA measures, which compare the individual team members’ communications to each other (Fusaroli et al., 2014; Fusaroli & Tylén, 2015). However, to the best of my knowledge only a single experiment and review paper have demonstrated this. Therefore, to evaluate and potentially bolster this claim, the following analytic technique was proposed partially based on the approach adopted by Fusaroli and Tylén (2015).

The analytic logic for differentiating patterns of team interaction dynamics, particularly between synchronized communicative behaviors and complementary communicative behaviors, warrants some further elaboration. On the one hand, if team members exhibit very similar team knowledge building behaviors and are synchronized in doing so (exhibiting those behaviors at or very close to the same time), then CRQA, comparing the two team members’ individual communications time-series, will be sensitive to this dynamic in the form of a high %REC, particularly around the line of coincidence (e.g., Fusaroli et al., 2014). On the other hand, if the team as a whole is exhibiting complementary communicative behaviors, then the %REC generated from RQA should be higher when compared to the %REC from the CRQA comparing
the two team members’ communications. For example, considering instances where there are adjacency pairs, such as a request for information followed by a provision of information, these would manifest in a recurrence plot (RP) of the whole team’s communications with twice as many recurrent points, when compared to a cross-recurrence plot (CRP) conducted on the two separate communications time-series representing each team member. That is, if a given team member is consistently requesting information and another team member is providing information, then the %REC will be low because they are exhibiting different behaviors. But, in taking the communications as a whole, instances where this form of complementary behavior is exhibited are manifested by higher recurrence, should this pattern of interaction exist.

A related notion is that if individual team members are contributing complementary behaviors, then the sequence of behaviors for a respective team member should center on a certain interaction sequence that is self-consistent and thus more recurrent with itself than it would be with the other team member. Figure 9 shows the a priori predicted pattern of findings necessary to support the claim that teams’ collaborative problem solving communications are characterized by complementarity as opposed to some other form of coordination such as synchronization.
Figure 9. Analytic logic to test for complementarity interaction pattern

(Plots made with Dale’s web-based recurrence plot generator:

Specifically, Figure 9 shows the prediction that %REC will be higher when computing RQA on the individual time-series of each team member (i.e., a form of self-consistency in behavior; see Fusaroli & Tylén, 2015) and on the team’s communications as a whole, when compared to %REC of individual team members’ time-series plotted against each other using CRQA. In Figure 8, P_# simply corresponds to a generic participant number for a given team.

In order to prepare the data to test this hypothesis, the following analyses were applied to the team communications time-series: a) RQA was applied to recurrence plots for the three time-series representing the teams’ overall communications (speaker, code, and code*speaker), b) CRQA was applied to the cross-recurrence plots of the four time-series representing each team members’ individual communicative contributions (i.e., speaker 1 crossed with speaker 2 and code speaker 1 crossed with code speaker 2), and c) RQA was applied to each individual’s
communicative contribution time-series (i.e., code speaker 1 and code speaker 2). For all of these analyses, each of the RQA measures shown in Table 6 were derived (i.e., %REC, %DET, MEANLINE, ENTROPY); however, only %REC for the code-based time-series are important for testing Hypothesis 3 and Hypothesis 4 (the rest are of import to answering the other research questions). Because all time-series were discrete and nominal, the parameters for phase space reconstruction used to generate the recurrence plots were as follows: a radius of .0001, an embedding dimension of $M = 1$, and a delay of $\tau = 1$ (Dale, Warlaumont, & Richardson, 2011). All recurrence quantification analyses were conducted using the CRQA package (Coco & Dale, 2014b) in R.

To test for the interaction pattern of complementarity, a series of paired-samples t-tests were conducted to make the %REC comparisons represented in Figure 8 for the code-based time-series only. As predicted, the %REC was significantly higher for the code time-series representing the collaborative problem solving communications of the whole team ($M = 13.01$, $SD = 1.85$) when compared to the %REC derived from comparing two team members code time-series with CRQA ($M = 2.89$, $SD = 0.52$), $t(42) = 42.81$, $p < .001$. In addition, the %REC for both the codespeaker1 ($M = 28.87$, $SD = 8.38$, $t(42) = 20.40$, $p < .001$) and codespeaker2 ($M = 24.14$, $SD = 1.15$, $t(42) = 138.22$, $p < .001$) time-series were significantly greater than the %REC derived from comparing the two team members’ code time-series with CRQA. Figure 10 shows an example observed comparisons for a single team in the form of the RPs and CRP predicted above. Note the similarity in the relative density of recurrent points between the a priori prediction and the observed recurrence pattern.
Given the analytic logic detailed here, and introduced by Fusaroli et al. (2014), the results demonstrated that teams exhibited a coordinated form of interaction in their collaborative communications that can be characterized as complementarity. This can be assumed again, because %REC was higher for the team communications taken as a whole and for each individual team members’ communication when compared to the cross-recurrence between team members. Additionally, the results lend support to idea that teams exhibited a dynamic pattern of communicative interaction that is better characterized by complementarity than by synchronization as predicted by the dialog as interpersonal synergy approach (Fusaroli et al., 2012; Fusaroli & Tylén, 2015).

Table 8
*Observed example of analytic logic illustrating evidence of complementarity interaction pattern.*

As a means of further illustrating this complementarity interaction pattern, two excerpts from transcriptions are shown below. The first example below shown in Table 9 is taken from a team who had a higher %REC indicating more complementarity. What is worth noting from this
example is the clear turn-taking dialog in a timely fashion with appropriate responses. When one team member, requests knowledge or delegates goal-direct tasks the other person typically provides an acknowledgment, the requested knowledge, or a clarification.

Table 9
*Communications transcription excerpt for high complementarity team.*

<table>
<thead>
<tr>
<th>Time Stamp</th>
<th>Speaker</th>
<th>Utterance</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>7:13</td>
<td>Right</td>
<td>Uh, for- to bring life support, we need the welding torch robot, right?</td>
<td>KR</td>
</tr>
<tr>
<td>7:19</td>
<td>Left</td>
<td>Um, yes.</td>
<td>S</td>
</tr>
<tr>
<td>7:21</td>
<td>Left</td>
<td>We'll also need something to grab and replace parts.</td>
<td>KP</td>
</tr>
<tr>
<td>7:23</td>
<td>Right</td>
<td>Alright, so then we're gonna need two robots.</td>
<td>KP</td>
</tr>
<tr>
<td>7:25</td>
<td>Left</td>
<td>Yeah.</td>
<td>S</td>
</tr>
<tr>
<td>7:34</td>
<td>Right</td>
<td>I'm gonna go and build the welding torch.</td>
<td>GTO</td>
</tr>
<tr>
<td>7:36</td>
<td>Left</td>
<td>Okay. Alright.</td>
<td>S</td>
</tr>
<tr>
<td>7:37</td>
<td>Right</td>
<td>Then I'll, uh, bring the rover over to go and bring it over.</td>
<td>GTO</td>
</tr>
<tr>
<td>7:46</td>
<td>Right</td>
<td>And it does look like, uh, jump running will actually speed you up a little.</td>
<td>KP</td>
</tr>
<tr>
<td>7:50</td>
<td>Left</td>
<td>Okay.</td>
<td>S</td>
</tr>
<tr>
<td>7:57</td>
<td>Left</td>
<td>Okay, we have power on the left system.</td>
<td>SU</td>
</tr>
<tr>
<td>8:05</td>
<td>Right</td>
<td>Uh, I cannot drive the robot.</td>
<td>SU</td>
</tr>
<tr>
<td>8:10</td>
<td>Left</td>
<td>You can't?</td>
<td>IR</td>
</tr>
<tr>
<td>8:11</td>
<td>Right</td>
<td>No.</td>
<td>S</td>
</tr>
<tr>
<td>8:14</td>
<td>Right</td>
<td>Oh, I can load it on the back, though.</td>
<td>KP</td>
</tr>
<tr>
<td>8:15</td>
<td>Left</td>
<td>Okay.</td>
<td>S</td>
</tr>
<tr>
<td>8:20</td>
<td>Right</td>
<td>You know, I can load two.</td>
<td>KP</td>
</tr>
<tr>
<td>8:21</td>
<td>Right</td>
<td>Do you want me to load the other one?</td>
<td>GTO</td>
</tr>
<tr>
<td>8:22</td>
<td>Left</td>
<td>Yeah, yeah.</td>
<td>S</td>
</tr>
<tr>
<td>8:24</td>
<td>Left</td>
<td>That way we can get them both into place.</td>
<td>SEval</td>
</tr>
</tbody>
</table>

Contrary to the example above, the communications excerpt shown below in Table 10 is for a team with low complementarity indicate by low %REC. This excerpt is characterized by a lack of timeliness and relevance of responses by the Right team member. If you notice there are instances where the Left team member will request information and the other team member does
not even respond. What is evident here is an asymmetry in the communicative contributions between the two team members. The Left team member seems to be providing relevant situation updates, knowledge, and goals, but rarely gets a response from the other team member, and if they do, it is often unrelated. Thus, this highlights a scenario in which the interaction is not as complementary as that described above.

Table 10
Communications transcription excerpt for low complementarity team.

<table>
<thead>
<tr>
<th>Time Stamp</th>
<th>Speaker</th>
<th>Utterance</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>7:27</td>
<td>Left</td>
<td>Are you fixing any or just?</td>
<td>SR</td>
</tr>
<tr>
<td>7:29</td>
<td>Right</td>
<td>Yeah, I have a wrench</td>
<td>SU</td>
</tr>
<tr>
<td>7:33</td>
<td>Left</td>
<td>You need a welder to uh, fix them.</td>
<td>KP</td>
</tr>
<tr>
<td>7:36</td>
<td>Right</td>
<td>This one says I need a wrench and it's not letting me,</td>
<td>IP</td>
</tr>
<tr>
<td>7:44</td>
<td>Left</td>
<td>Wow, it's a lot to fix.</td>
<td>KP</td>
</tr>
<tr>
<td>7:56</td>
<td>Left</td>
<td>Darn it.</td>
<td>INC/F/EX</td>
</tr>
<tr>
<td>8:01</td>
<td>Right</td>
<td>You said Tonya?</td>
<td>IR</td>
</tr>
<tr>
<td>8:02</td>
<td>Left</td>
<td>No, I said darn it.</td>
<td>IP</td>
</tr>
<tr>
<td>8:03</td>
<td>Right</td>
<td>Oh, okay.</td>
<td>S</td>
</tr>
<tr>
<td>8:05</td>
<td>Left</td>
<td>Alright it says I repaired it but it's still red so I don’t understand that um</td>
<td>UNC</td>
</tr>
<tr>
<td>8:11</td>
<td>Left</td>
<td>Any idea or no?</td>
<td>KR</td>
</tr>
<tr>
<td>8:21</td>
<td>Left</td>
<td>Hey, don’t go over by the cooler, over where I just was.</td>
<td>GTO</td>
</tr>
<tr>
<td>8:25</td>
<td>Left</td>
<td>You need a robot for that.</td>
<td>KP</td>
</tr>
<tr>
<td>8:28</td>
<td>Right</td>
<td>Oh.</td>
<td>INC/F/EX</td>
</tr>
<tr>
<td>8:40</td>
<td>Left</td>
<td>Um</td>
<td>INC/F/EX</td>
</tr>
<tr>
<td>8:42</td>
<td>Right</td>
<td>(laughs)</td>
<td>INC/F/EX</td>
</tr>
<tr>
<td>9:01</td>
<td>Left</td>
<td>Do you know where to get a robot or no?</td>
<td>KR</td>
</tr>
<tr>
<td>9:03</td>
<td>Left</td>
<td>Okay, I'll have to find it.</td>
<td>GTO</td>
</tr>
<tr>
<td>9:06</td>
<td>Right</td>
<td>Do you want me to find one?</td>
<td>GTO</td>
</tr>
<tr>
<td>9:08</td>
<td>Left</td>
<td>Um</td>
<td>INC/F/EX</td>
</tr>
<tr>
<td>9:10</td>
<td>Left</td>
<td>I just can't really remember which</td>
<td>UNC</td>
</tr>
<tr>
<td>9:13</td>
<td>Left</td>
<td>which pl…, uh, which storage thing it was in.</td>
<td>KR</td>
</tr>
<tr>
<td>9:16</td>
<td>Left</td>
<td>Just head over to that one and see If you can find one.</td>
<td>GTO</td>
</tr>
<tr>
<td>9:26</td>
<td>Left</td>
<td>Alright, repaired that.</td>
<td>SU</td>
</tr>
<tr>
<td>9:32</td>
<td>Left</td>
<td>I'm going to repair all the couplers.</td>
<td>GTO</td>
</tr>
<tr>
<td>9:45</td>
<td>Right</td>
<td>Sorry I didn’t mean to do that I'm just trying to get out of the chatbox.</td>
<td>T/OT</td>
</tr>
</tbody>
</table>
**Testing Hypothesis 4: Does complementarity in team communications predict better problem solving outcomes?**

Since the results of Hypothesis 3 supported the idea that teams exhibited a complementary interaction dynamic in their problem solving communications, the key measure suggesting this, %REC for the code time-series of the whole team communications, is used for the following analyses.

A hierarchical linear multiple regression model was used to examine the relationship between teams’ communicative complementarity and problem solving performance while controlling for video game familiarity, knowledge assessment scores, and cognitive load using SPSS version 21.

MinVidGame, MinKnow, and AveCogLoad, were included as the control variables and thus entered into Step 1 of the model. The predicted variable was performance on the Moonbase Alpha collaborative problem solving task. Just as in the prior analysis (see Hypothesis 2 results), the control variables MinVidGame, MinKnow, and AveCogLoad accounted for a significant 53% of the variance ($R^2 = .53, R^2_{adj} = .50$), $F(3, 39) = 14.89, p < .0005$. All control variables were uniquely significant predictors of performance.

In Step 2, complementarity, operationalized as the %REC for the teams’ code communications time-series was added to the model as a predictor of performance. The addition of complementarity as a predictor added a non-significant percentage of variance to the model; less than 1%, $\Delta R^2 = .003, F_{\Delta R^2}(1, 38) = .221, p = .64$. Taken together, complementarity, MinVidGame, MinKnow, and AveCogLoad accounted for 54% of the variance in performance, ($R^2 = .54, R^2_{adj} = .49$), $F(4, 38) = 10.99, p < .0005$. In this model, again all control variables were
uniquely significant predictors of performance. However contrary to Hypothesis 4, complementarity, at least as characterized by the %REC for the code time-series, was not a uniquely significant predictor of performance, ($r^2 = .01, t(38) = .47, p = .64$).

Table 11 summarizes the results of this hierarchical multiple linear regression analysis including the standardized coefficients. As with the prior regression analysis, assumptions of the model were assessed in two ways. First, visual inspection of the residual plots and histogram did not reveal any clear deviations from homoscedasticity or normality. Also, results from the gvlma function in R showed that the assumptions of the model were not violated, thus suggesting it was an appropriate model of the data. Thus, the current results did not show support for H4 indicating that complementarity in teams’ collaborative problem solving communications did not predict performance on the Moonbase Alpha collaborative problem solving task.

Table 11
*Complementarity predicting performance multiple regression output.*

<table>
<thead>
<tr>
<th>Step</th>
<th>Variable</th>
<th>B</th>
<th>SE B</th>
<th>β</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Constant</td>
<td>-48.44</td>
<td>35.77</td>
<td>-1.35</td>
<td>.18</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MinVidGame</td>
<td>14.01</td>
<td>3.80</td>
<td>.410</td>
<td>3.68</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>MinKnow</td>
<td>10.11</td>
<td>3.46</td>
<td>.323</td>
<td>2.92</td>
<td>.006</td>
</tr>
<tr>
<td></td>
<td>AveCogLoad</td>
<td>-9.65</td>
<td>2.67</td>
<td>-.399</td>
<td>-3.61</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>Constant</td>
<td>-50.20</td>
<td>35.78</td>
<td>-1.40</td>
<td>.169</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MinVidGame</td>
<td>14.56</td>
<td>3.84</td>
<td>.426</td>
<td>3.79</td>
<td>.001</td>
</tr>
<tr>
<td>2</td>
<td>MinKnow</td>
<td>9.40</td>
<td>3.53</td>
<td>.300</td>
<td>2.66</td>
<td>.011</td>
</tr>
<tr>
<td></td>
<td>AveCogLoad</td>
<td>-9.67</td>
<td>2.67</td>
<td>-.400</td>
<td>-3.62</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>Complementarity</td>
<td>.907</td>
<td>1.93</td>
<td>.053</td>
<td>.470</td>
<td>.641</td>
</tr>
</tbody>
</table>
Testing Hypothesis 5: Does higher synchronization in bodily movement and complementarity in team communications predict better problem solving outcomes, when compared to other groups?

In order to test Hypothesis 5, an unweighted effects coding approach to linear multiple regression analysis was applied (Cohen, Cohen, West, & Aiken, 2003). First, however, high versus low values of synchronization and complementarity were determined by applying a median-split to the observed cross-correlation values and %REC values, respectively. Those above or equal to the median were assigned to the high group for each interaction measures and those below were assigned to the low group. The pairing of these two values contributes to a total set of theoretical combinations forming four groups (see Table 9).

Table 11
*Theoretical combinations of high vs. low synchronization and complementarity.*

<table>
<thead>
<tr>
<th>Synchronization</th>
<th>Complementarity</th>
<th>HSHC: High Synch/ High Comp</th>
<th>HSLC: High Synch/Low Comp</th>
<th>LSHC: Low Synch/ High Comp</th>
<th>LSLC: Low Synch/Low Comp</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>High</td>
<td>HSHC: High Synch/ High Comp</td>
<td>HSLC: High Synch/Low Comp</td>
<td>LSHC: Low Synch/ High Comp</td>
<td>LSLC: Low Synch/Low Comp</td>
</tr>
<tr>
<td>Low</td>
<td>Low</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For the present purposes, the group of greatest interest is those teams who are both high synchronization and high complementarity. However, four groups were created to represent these possible combinations. The number of observations in each group were as follows: HSHC = 13, HSLC = 11, LSHC = 11, LSLC = 10. Following the guidelines provided by Cohen et al. (2003), three effects coding variables were developed with HSLC as the base group.

Specifically, multiple linear regression was conducted with problem solving performance as the predicted variable and with three effects coding variables representing different
combinations of high versus low grouping of synchronization and complementarity. Results showed that this model accounted for a non-significant 9% of the variance ($R^2 = .09$, $R^2_{adj} = .02$), $F(3, 39) = 1.32, p = .281$. In directly testing Hypothesis 5, the results did not suggest that the coding variable, representing teams exhibiting both high synchronization and high complementarity, predicted a mean performance outcome significantly higher than the average mean across the other groups. Thus, these results can be interpreted as not supporting H5 because teams who exhibited higher degrees of synchronization and higher communicative complementarity were not more likely to perform better than teams who may have had heterogeneous pairings of these two interaction dynamics or that received low values on both.

Research Questions 1 & 2: Which communications time-series and which RQA-based measures are most sensitive to differences in problem solving performance?

In order to test the two research questions, three exploratory stepwise multiple regression models were computed for each respective team communications time-series (speaker, code, and code*speaker) with each of their respective recurrence-based metrics (%REC, %DET, MEANLINE, ENTROPY) as predictors and collaborative problem solving performance as the predicted variable. Due to the lack of findings relating the recurrence-based measure of complementarity in the prior analyses, only teams whose performance was greater than 0 were included here ($N = 32$). Entry criteria was $p < .099$ and removal criteria was $p > .10$.

Specifically, for the speaker and code*speaker time-series, none of the recurrence-based predictors met the entry criteria. However, two recurrence-based predictors from the code time-series did meet the entry criteria: %REC and ENTROPY. ENTROPY was entered in the first step, and %REC was entered in the second step. Results from the final model are shown in Table 10. This model accounted for a near significant 43% of the variance in performance ($R^2 = .43$,
$R^2_{adj} = .19$, $F(2, 29) = 3.31, p = .051$. In light of the proposed research questions, it seems that the *code* communications time-series and the %REC and ENTROPY RQA-based measures associated with this time-series were more sensitive to differences in problem solving performance than any of the others, at least for this particular task. Note, however, that this is only when excluding participants who received zeros for performance from the analysis, so this finding is only exploratory and tenuous.

Table 12
*Regression results from three communications time-series with RQA-based measures.*

<table>
<thead>
<tr>
<th>Communication Time-Series</th>
<th>Variable</th>
<th>B</th>
<th>SE B</th>
<th>β</th>
<th>t</th>
<th>p</th>
<th>$R^2_{adj}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constant</td>
<td>-73.85</td>
<td>45.66</td>
<td>-1.62</td>
<td>.827</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Code</td>
<td>%REC</td>
<td>5.32</td>
<td>3.00</td>
<td>.297</td>
<td>1.78</td>
<td>.086</td>
<td>.19</td>
</tr>
<tr>
<td></td>
<td>ENTROPY</td>
<td>114.85</td>
<td>61.83</td>
<td>.311</td>
<td>1.86</td>
<td>.074</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER FIVE: DISCUSSION

The present research sought to investigate the effects that certain forms of team interaction dynamics have on performance during a complex collaborative problem solving task. Based on extant work in the study of social interactions, I predicted that teams would exhibit interpersonal synchronization during the task that was at a level greater than chance (H1). Results from the present research supported H1, in that teams’ demonstrated a pattern of interaction in their bodily movements that was characterized as synchronized. Moreover, this was significantly different from a baseline level of synchronization generated using a shuffled surrogate analysis. These findings contribute to research on problem solving because they provide evidence that synchronization is an emergent phenomena that occurs during a complex, collaborative problem solving task that is heavily mediated by technology (i.e., the use of a computer). In particular, much of the prior work finding evidence of synchronization during social interaction was on simpler tasks and allowed for participants to have a greater degree of freedom of movement during their tasks (e.g., Paxton & Dale, 2012). Thus, this finding illustrates that even during a complex task that required the use of mouse and keyboard, teams exhibited a highly coordinated pattern of bodily movement in which, over the duration of the task, they were moving at approximately the same time.

Contrarily, the results did not lend support to H2. That is, when controlling for video game experience, task knowledge, and cognitive load, interpersonal synchronization did not have a significant relationship with collaborative problem solving performance. Teams who performed poorly seemed equally likely to exhibit the same amount of synchronization in their bodily movement as those teams who performed well. This is somewhat surprising given that
Paxton et al. (2013) found that performance on a problem solving task was predicted by synchronization of bodily movements. However, Abney, Paxton, Dale, and Kello (2015) found that it was actually lower degrees of synchronization that were predictive of better problem solving performance. The present research did not observe either relationship. Note, though, that prior research did not control for factors such as cognitive load (or relevant experience factors). As such, it is unclear whether earlier studies would have shown a relationship between synchrony and performance when additional control factors were included.

While a number of independent research projects have found the emergent pattern of synchronization across a variety of tasks, this is the first research of its type to examine the effects of synchronization in a complex, technology-mediated collaborative problem solving task. Related research has examined synchronization of various modalities such as specific bodily movements (Louwerse et al., 2012), postural sway (Shockley et al., 2003), and heart-beat intervals (Strang et al., 2014). One explanation for lack of support for H2 is that the gross level of bodily movements captured by the frame-differencing method may not provide the necessary level of specificity. For example, a behavioral coding scheme such as that adopted by Louwerse et al. (2012) might detail which particular behavioral modalities become synchronized and the relationship they have with performance (e.g., posture, facial expression, etc.). It could be that the synchronization of a particular behavioral modality is predictive of performance.

Another potential explanation is that the task itself may require a certain amount of bodily synchronization that does not necessarily relate systematically to performance. One example of this is provided by Strang et al. (2014) who tested for the relative effects of the task constraints on their observation of specific coordination dynamics using a virtual-pairs analytic
technique. This seems plausible given that participants need to exhibit some consistent movements related to their keyboard strokes and mouse movements, while there are more degrees of flexibility related to their bodily movements such as postural adjustments. One way to examine this would be by recording and analyzing the synchronization between participants’ mouse movements (e.g., Dale et al., 2011). Additionally, much of the prior research showing interactional benefits from synchronization actually required participants to engage in deliberately synchronized behaviors and then their performance on some subsequent task was improved (e.g., Hove & Risen, 2009; Valdesolo et al., 2010). It could be that to observe a benefit of synchronization, such as better problem solving performance, participants may need some explicit awareness of the synchronization of their movements.

Another explanation is that in such a complex collaborative problem solving task, the role of bodily movements and the interaction between teammates may not be paramount. Instead, as much of team cognition research predicts, it could perhaps be that the amount of similarity or overlap of shared knowledge structures of the team is more crucial (e.g., DeChurch & Mesmer-Magnus, 2010), at least when compared to synchronization of bodily movements. Regardless of its relationship with performance, synchronization of bodily movements was prevalent in this experiment beyond a level that could be attributed to chance. Therefore, the effects of synchronization on CPS performance should be studied using other knowledge intensive CPS tasks while also examining team shared knowledge structures. In short, more scrutiny of the coupling of cognitive processes (knowledge in the head) and behavioral processes (communication and synchrony) may illuminate associations occurring in complex collaborative tasks.
As predicted by H3, teams did exhibit a form of collaborative problem solving communications that could be characterized as complementary. This was evidenced by a higher %REC for the code time-series representing the whole team when compared to the cross-comparison of team members. The cross comparison was also lower than the self-consistent pattern evidenced in individual team members’ communications. The idea here is that if team members are engaging in complementary behaviors then each should have a consistent pattern of behavior that is distinct and that the overall team communication will have higher recurring patterns given these distinct and complementary behaviors of each team member. In addition, example excerpts were provided that highlighted a case of high and low complementarity from communications transcriptions. These showed that a team with high complementarity exhibited clear turn-taking dialog in a timely fashion with appropriate responses; whereas, the team with low complementarity exhibited an asymmetric interaction pattern in which one team member provided task relevant communications, only to either not receive a response or an irrelevant response in most cases.

One issue with this interpretation is the degree to which the %REC of code communications time-series actually represent the signature of complementarity. At least according to the logic here, developed in part by Fusaroli and Tylén (2015), in which comparisons were made between cross-RQA (recurring communication behaviors with each other) and individual RQA representing self-consistent patterns (recurring communication patterns of the individual), this seems a sound interpretation. However, some simulation studies might be able to further differentiate and test this logic to a greater extent. Nonetheless, this interpretation is justified based on the extant, albeit limited, research using RQA in this manner. Some alternative metrics to consider would be, rather than using a single measure of
complementarity, instead test the fit of regression model between the recurrence measures of the dyad’s dialog and the cross-recurrence between the dyad member’s individual dialog (Fusaroli & Tylén, 2015). They found the model that included the complementarity of the interaction was the better fit and predictor of task performance. Alternatively, Sadler et al. (2009) used cross-spectral coherence as a metric of complementarity for interpersonal dimensions during an interaction. It could be possible to apply this technique to other aspects of communications, but it likely would not apply to nominal communications time-series. Another option could be to use orbital decomposition analysis, which has previously been applied in group problem solving contexts to determine the prevalence of ideal complementary sequences for each team (Guastello, 1998; Guastello, Hyde, & Odak, 1998) and test the relationship between the prevalence and performance.

In examining H4, the results did not show a significant positive relationship between complementarity, indicated by %REC of the code time-series, and collaborative problem solving performance on the Moonbase Alpha Task. In other words, it did not seem that the complementary communicative interaction patterns were predictive of better problem solving performance. This is also surprising given the theoretical underpinnings and findings in prior research (e.g., Fusaroli et al. 2012). Perhaps other more foundational linguistic aspects of dialog matter more in this context such as prosody and speech/pause dynamics (Fusaroli & Tylén, 2015) as opposed to recurrence of semantic content. However, given the plethora of research demonstrating that semantic content and particular sequences of communicative acts are related to team performance (e.g., Bowers et al., 1998, Fischer, McDonnell, & Orasanu, 2007), this does not seem likely. It could be that this recurrence-based measure is at too high a level of granularity (i.e., a sort of indiscriminate recurrence of certain forms of communication), that it
does not relate systematically to performance. Potential methods to address this are discussed in the Limitations and Future Directions section.

Given the lack of the systematic relationship between interpersonal bodily movement synchronization and communicative complementarity and performance, no support was found for H5. That is, the test for H5 examined whether there was a relationship between groupings of these forms of interaction into high and low categories of synchronization and complementarity with performance. In particular, it was predicted that high synchronization and high complementarity would be predictive of better performance compared to other groupings such as low synchronization and low complementarity. The results did not support this hypothesis.

Lastly, the present research also sought to determine which type of team communications time-series and which recurrence-based measures would be most sensitive to differences in performance. The results showed that the code communications time-series and the ENTROPY and %REC recurrence-based measures were more sensitive to differences in performance. This interpretation was arrived at based on the fact that none of the other time-series and their associated measures met the entry criteria to be considered as predictors of performance. Given the positive beta values for this model, it could be inferred that teams who exhibit collaborative problem solving communications that recur more frequently (higher %REC) and in different recurring sequences of different lengths (higher ENTROPY) are likely to have better problem solving performance. However, the code time-series model was only nearly significant in predicting performance ($p = .051$). This could perhaps be due to the smaller sample size resulting from excluding participants who did not score higher than zero on performance from this analysis. While the application of recurrence quantification analysis has only been applied in
a handful of team communications studies, prior work has found success in relating recurrence-based measures from communications data to performance (e.g., Gorman et al., 2012a; Russell et al., 2012), thus this lack of significant relationships for any of the time-series and for any of the measures was surprising. However, the exploratory analyses adopted here could be improved in future work using stronger theoretical predictions indicating good ranges for the RQA-based metrics. For example, Gorman et al. (2012a) used a three-state, speaker only time-series and generated cases of complete determinism, randomness, and combined determinism and randomness for various time-series lengths. They argued that effective and adaptive team performance should have %DET values that are in between completely deterministic and completely random. Future work should attempt to make such predictions regarding the other RQA-based metrics in the context of team communication during CPS.

In summary, the results showed that teams did tend to synchronize their bodily movements at a level significantly greater than chance and that their communications could be characterized as complementary. The results did not, however, show that either of these interaction measures were predictive of problem solving performance. Further, examination of the different team-communications time-series and their corresponding recurrence-based measures suggested the code time-series conveying semantic content and the associated %REC and ENTROPY values may be most sensitive to performance, but due to the lack of significant findings, these interpretations should be received with caution and warrant further examination. Additionally, the results did show that video game experience, task knowledge, and perceived cognitive load were significant covariates that were predictive of problem solving performance.
Theoretical and Practical Implications

In a general sense, this dissertation contributes to basic science through advancement of theory and the combination of previously unstudied factors that play a role in CPS. It aimed to advance practical application through examination of potential unobtrusive measures for predicting good and poor team process during CPS. Overall, this dissertation adds scientific knowledge by advancing theory and empirical knowledge on the emergence of different forms of verbal and non-verbal interaction patterns during CPS, but says less about what forms of interaction are predictive of effective problem solving performance.

With regard to the Macrocognition in Teams theoretical and empirical approach, and CPS more generally, this research is the first to empirically examine and utilize analytic methods that capture and quantify the dynamic nature of team interaction during collaborative problem solving. Whereas much of the prior research has only investigated retrospective reports of collaborative problem solving (e.g., Fiore et al., 2014) or linearly predictive accounts of how specific problem solving communications are related to performance (Rosen, 2010), the approach adopted by the present research utilized measures that persevered the temporal dynamics of the team interaction and used those to predict performance. The methods employed here are consistent with original conceptualizations and terminology espoused in theory building efforts related to Macrocognition in Teams (Fiore et al., 2010a; 2010b). Even though the interaction-based measures were not significantly related to CPS performance, interaction-based measures, both verbal and non-verbal, should still be viewed as necessary factors for understanding collaborative problem solving. By contrast, it would seem preposterous to draw a conclusion that interaction does not matter during CPS as, by definition, collaboration is an interactive process.
With regards to theoretical and empirical work investigating interaction or coordination dynamics as well as approaches examining dialog as interpersonal synergy, this research provides compelling implications. For example, to the best of my knowledge this is the first experiment to find evidence of interpersonal synchronization of bodily-movements during a complex, technology-mediated collaborative problem solving task that was very knowledge intensive. While other tasks have may share similar aspects such as the map-task (e.g., Louwerse et al., 2012), which is also computer-mediated, the present task was very complex and difficult for participants, relatively speaking. Not only does this research add to the robust findings of interpersonal synchronization across a wide variety of tasks, but it also provides evidence for complementary communication dynamics during CPS, as would be predicted by the dialog as interpersonal synergy approach. To date, relatively little empirical work has provided evidence of this interaction pattern (i.e., Fusaroli & Tylén, 2015) with these methods. In short, the primary implication is additional evidence for these theoretical approaches and the utility of dynamical systems methods for identifying particular forms of team interaction that emerge during CPS.

With regard to team cognition, the theoretical implication of the present work is emphasis on a more consistent theoretical mapping of team cognition as interaction (Cooke et al., 2012) with the appropriate research methods to identify dynamic interaction patterns. However, in interactive team cognition theory, measures of team interaction are argued to be able to account for more variance in performance than knowledge-based metrics. In addition, Cooke et al. (2012) point out that knowledge based measures may be well suited for explaining performance in recently formed, small and homogeneous teams working on knowledge-oriented tasks. This description could certainly characterize the teams in the present study. Given that the findings of the present work did not demonstrate the relationship between team interaction and performance,
the implication here is that interaction and shared cognition approaches should be co-investigated in future work to provide a richer explanation of team cognition during CPS.

In terms of practical implications, the present work demonstrates methods that can be used to observe and quantify interaction dynamics during CPS in complex work domains. Specifically, this study used the video frame-differencing method, which had only been applied to studies of conversations and simple problem solving tasks. Here it was found to be an objective and relatively time-efficient way to capture bodily movement synchronization compared to traditional methods (e.g., Bernieri & Rosenthal, 1991). Additionally, this study demonstrated the use of recurrence quantification-based measures derived from team communications, which only a handful of prior studies have done. This work, in particular, serves as additional practical guidance for how to apply these techniques to identify particular forms of verbal and non-verbal team interaction, but it does not provide insights as to what measures are likely to be most sensitive to differentiating performance outcomes.

In other words, the primary practical implication here is demonstration of the application of methodological tools to study CPS that are consistent with prior theoretical descriptions of CPS processes and team interaction. Because the current tools were successful in identifying particular forms of interaction, but these forms were not systematically related to performance, further research is needed in applying these methodological tools, and possibly other tools, to better understand effective CPS performance. Potential ways of improving the application of these methodological tools and some others to consider in future work are detailed in the next section.
Limitations and Future Directions

The present research has a number of limitations that, to some degree, warrant a conservative interpretation of the findings. These include issues related to: the laboratory setting and college sample, the nature of the task, the way CPS performance was operationalized, the time intensiveness of the communications coding process, the lack of specificity of the communications analysis, the potential existence of other forms of coordinated interaction, and critique of whether a dyad is really a team. Each of these issues is discussed in more detail and potential future research directions to address these issues are proposed.

One limitation of the present research is that it was conducted in a laboratory setting using a college undergraduate student population. The Macrocognition in Teams theoretical framework was developed to advance scientific understanding of how knowledgeable professionals are able to collaboratively solve problems in operational, complex tasks. The fact that there has been empirical evidence for these collaborative problem solving processes from work settings (Hutchins, 2010; Fiore et al., 2014) and laboratory tasks with college populations (Rosen, 2010; Seeber et al., 2013) as well as the reliability of the coding process presented here, does bolster the ecological validity of the present findings regarding the communications data, however tentative they may be. Future research could attempt to replicate this study in contexts with experienced professionals working in operational contexts, but given the costs that could be associated with this, further laboratory research is needed as a first step to better understand the relationship between interaction and performance.

Another limitation of this study is the nature of the task that participants completed. NASA’s Moonbase Alpha task (NASA, 2011) was selected due to the features of the task that
qualify it as complex and necessitating both collaboration and problem solving (see Method section). Of course, this represents just one of many potential tasks in which complex collaborative problem solving is required. Collaborative problem solving tasks may vary greatly in the number and type of factors that qualify them as complex. Given the somewhat unique combinations of factors and variables involved in the task, the results of this study could be unique to this particular task. Future work should, more generally, see if similar findings are evident when examining team interaction dynamics across a variety of collaborative problem solving tasks. If the findings are different, steps should be taken to develop comparability across collaborative problem solving tasks by identifying the factors and variables that are present for a given task. One way of doing so might be to employ methods for quantifying task complexity more systematically (e.g., Wood, 1986) or other task analysis measures (e.g., Arthur, Edwards, Bell, Villado, & Bennett, 2005).

An important limitation is the way that CPS performance was operationalized in the present study. Recall that a combined performance measure was developed that accounted for the total time to complete the task and the amount of oxygen restored. Given the nature of the task and the goals (i.e., to restore 100% oxygen in the least amount of time possible), this measure makes sense. However, given repeated observation of task performance, an issue became evident. Specifically, zero values of performance could reflect a variety of things. From observations of participants performing the task, it seemed that the lack of executing a particular action within the simulation could result in participants completing many of the repairs without restoring any oxygen. Specifically, if participants did not secure the power cables to the power couplers with a wrench after attaching them, oxygen restoration would not commence. This is
problematic, because these teams could be very highly coordinated teams who are lacking a particular piece of task knowledge and they would receive the same score as teams who were exceedingly uncoordinated and perhaps, not even taking the task seriously. Future research efforts could involve a performance measure with more variability than the one included here. On the one hand, a different way to measure performance would need to be determined, for example, by measuring the total object repairs a team makes; or, at a minimum, whether or not teams did recognize they had to secure the power cables with the wrench. This of course, could take an extensive amount of post-hoc examination of the video playback of teams’ performance, but, in turn, it might provide a more nuanced and valid measure of performance that relates systematically to the interaction-based measures.

In studies of team cognition and coordination, two primary approaches currently exist with different explanations regarding the important contributors to effective team performance: shared knowledge approaches and dynamical approaches (see Cooke et al., 2012; Gorman, 2014). The present research would fall under the dynamical approach and thus the formulated hypotheses and questions are aligned with that approach (i.e., a focus on the dynamics of team interaction and the effects certain patterns of interaction have on team performance). However, as Gorman (2014) points out, a general theory of team coordination and its relationship with effective task performance is still needed. Given the findings related to task knowledge and cognitive load, future research could examine the relationship between shared team knowledge structures and team interaction dynamics and the relative effects each measure contributes to effective team performance.
Another opportunity is an improvement in methodology related to examining team communications. On the one hand, the coding method developed here was very rigorous, but that comes at a great temporal cost. Such methods, may be of little utility to practitioners desiring to understand team performance in, or near, real-time (Gorman, Hessler, Amazeen, Cooke, & Shope, 2012b). One method for future research to consider applying to team communications during CPS would be conceptual recurrence, which integrates methods from computational linguistics and recurrence quantification analysis in order to provide an automated technique to understand the conceptual patterns that recur during human communication and interaction (Angus, Smith, & Wiles, 2012a; Angus, Smith, & Wiles, 2012b). Future research could corroborate findings using conceptual recurrence with findings based on traditional coding schemes. If there was any similarity with the two methods, this could provide a more automated and time-efficient method for understanding team communication during CPS. While not providing a way to save time, alternate coding schemes could be applied to team communications data to derive additional insights. For example, Fischer et al. (2007) applied a multi-level coding scheme to examine team communications that categorized interactive and affect-related statements in addition to task-related factors (such as problem solving). Exploring ways to automate communications coding techniques will afford the opportunity for real-time evaluation of team communications along a number of dimensions and is thus, crucial to future research. In addition, the more nuanced a coding system can be, the better understanding of CPS may be ascertained. Of course, these two dimensions are to some degree at odds, but nonetheless, represent practical consideration for researchers in this area.
An additional issue related to communications coding and the analyses of the present work is the degree to which it provides insights regarding the specific types of collaborative problem solving processes, the frequency in which they should occur, and in what sequences. In fact, the current analytic techniques, at least as they were applied here, do not provide such insights at all. Indeed, analyzing the frequency of team communications (e.g., Volpe, Cannon-Bowers, Salas, & Spector, 1996) or the optimal sequences that certain forms of communications should take (e.g., Bowers et al., 1998) have already shown to be effective indicators of team performance, more generally. Further, Rosen (2010) applied such techniques to understanding communication during collaborative problem solving, more specifically. In the present study, recurrence quantification measures were applied to the teams’ communications in order to move beyond traditional methods for examining communications in order to identify a particular dynamic pattern of interaction, complementarity. As Russell et al. (2012) noted, recurrence-based measures are informative and complementary to the more traditional frequency and sequence communication measures. Thus, future work could compare the frequencies of certain communications codes between high and low performing teams and apply analytic techniques such as lag-sequential or multi-way frequency analyses to identify which communication sequences are associated with better performance (however, see Dale et al., 2011 for how cross-recurrence analysis is related to lag-sequential analysis). Another option for future work that would be more aligned with applying dynamical systems techniques to understand CPS is to identify the optimal length and sequences of CPS communications using orbital decomposition analysis, which has previously been applied to group problem solving contexts (Guastello, 1998; Guastello et al., 1998).
Synchronization and complementarity are two forms or patterns of team coordination dynamics that prior theory and empirical work supported investigating in the present work. However, in terms of interaction and coordination dynamics, these two are not an exhaustive list of patterns and thus, others are worth examining. For example, a recent model of coordination developed by Butner, Berg, Baucom, and Weibe (2014) allows for further specification of different forms of synchronization. Specifically, Butner et al. (2014) use latent-change score modeling to determine whether a system is synchronized in a time-locked fashion and if so, whether it is in-phase or anti-phase, whether a system is entrained (there is a systematic relationship characterized by periods of synchrony and periods of desynchrony), or whether there is no synchronization prevalent at all. Relatedly, Abney and colleagues (Abney et al., 2014; Paxton et al., 2014) have looked at patterns of behavioral matching and complexity matching. On the one hand, behavior matching is similar to the methods employed here examining overt behaviors. On the other, complexity matching examines the distributional properties of interaction variables to examine patterns of coordination that exist across scales of measurement (see also Marmelat & Delignières, 2012). As another example, Strang et al. (2014) examined the coupling of various behavioral and physiological data in order to identify the complexity and regularity of those patterns. Further, Gorman et al. (2012b) focus specifically on the stability of the coordination as well as identifying parameters that control and order the coordination dynamics (Gorman et al., 2010). The point here is that many different forms of interaction and coordination dynamics exist that can be identified using a variety of analytical techniques and thus, there are many potential directions for future work.
In addition to identifying patterns of coordinated interaction in future work, researchers may be concerned with the causality of interaction patterns and their relationship with team performance (i.e., does synchronization lead to good performance or does good performance lead to synchronization). To some degree, this it at odds with dynamical approaches, which typically describe an emphasis on multi-scale emergent self-organization of the system being studied. Attempts at isolating components, and establishing causality, have been argued to be impossible (e.g., Richardson et al., 2014). Indeed, in the context of the present CPS research, task performance and the interaction patterns emerged and co-occurred across the duration of the task, making it difficult to establish temporal precedence for one or the other. Likewise, in interactive team cognition theory, a reciprocal causality is described where team-level cognition constrains interaction and individual-level contributions constrain team level interaction (Cooke et al., 2012). Prior work has established temporal precedence of synchronization by explicitly requiring dyads to synchronize and then examining performance on a subsequent task (e.g., Valdesolo et al., 2010). Although this might be a future direction for establishing a causal relationship between interaction patterns and task performance, it is arguably a different sort of interaction pattern because it is not emergent during the task itself. Furthermore, it may be the case that understanding causality requires additional, and more sophisticated, measures capable of better integrating interactions, communication, and cognition. Current methods are limited by post-hoc integration of these facets of team cognition, thus potentially limiting the possibility of uncovering more complex patterns of causality.

A criticism of operationalizing this work as investigating team interaction dynamics is that while a team was defined as two or more individuals working towards a common goal (Salas
et al., 1992), some have argued a team or group cannot be a dyad and must include three or more individuals (Moreland, 2010). The present analytic methodologies would be difficult to utilize in teams of three or greater, although similar techniques could be adopted and should be pursued in future research. For example, RQA could be applied on teams’ communication (e.g., Gorman et al., 2012a), but CRQA could not be applied to compare more than two team members at a time (Fusaroli et al. 2014). Therefore, the logic for identifying complementarity used here would not be able to be used for teams of three or greater. However, conceptual recurrence could be applied to multiple team members’ communication time series (Angus et al., 2012a; 2012b) and it may be possible to derive a metric of conceptual complementarity. Additionally, joint recurrence plots and quantification could be a potential way to derive a recurrence plot that represents the cross comparisons needed for examining complementarity, although such a method has not been conducted before (cf. Arcentales, Giraldo, Caminal, Benito & Voss, 2011 for use of joint recurrence across different behavioral signals). Further, Butner et al.’s (2014) latent change score modeling method would allow for examining various forms of synchronization in more than two individuals. And lastly, Duarte et al. (2013) developed a cluster phase analytic technique for looking at how large teams of European football players synchronized their movements. In short, there are a variety of related techniques that could be applied and adapted to examine the types of interaction dynamics investigated in the present work for teams greater than two individuals. It seems there are relatively clear ways to examine synchronization in teams with more than two individuals, but complementarity might be more difficulty. Future efforts should certainly develop such methods to investigate team interaction dynamics in larger teams.
Lastly, in the CPS literature, there is often discussion of phases of problem solving (Bales & Strodtbeck, 1951; Fiore et al., 2010a; 2010b). Two items related to this are at issue. The first is whether gross-level measures of bodily synchronization and/or communicative complementarity are appropriate in this context. This is an issue because, while the methods employed here preserve temporal relationships, they still provide a characterization of the interaction across the entire duration of the CPS task. To illustrate why this could be problematic, it could be that, during the initial phases of problem solving, synchronization or complementarity may matter less because teams are contributing unique information as they try to map out the problem space (Fiore & Schooler, 2004) and construct knowledge (Fiore et al., 2010a). But, during later phases, such as evaluation and revision, complementarity may become more important because team members must focus on evaluating problem solutions, discussing positive and negative aspects of the solution, and generating alternatives. This would show up as a specific form of complementarity, but require a precise level of granularity that is phase specific. Relatedly, the second issue is that effective teams should iterate through theoretically predicted problem solving phases (see Bales & Strodtbeck, 1951; Fiore et al., 2010a; 2010b); however, the present work did not attempt to identify such phases. Recent dynamical systems methods have been developed to identify phase transitions during problem solving (Stephen, Boncoddo, Magnuson, & Dixon, 2009), and these could be adopted for analysis of team communications. Upon identification of distinct problem solving phases, the metrics of synchronization and complementarity could be applied to the specific aspects of time-series of particular phases. It may be that identification of these interaction patterns during particular phases of problem solving is more important for predicting performance than the gross-level measures employed in this research.
CHAPTER SIX: CONCLUSION

This dissertation has provided evidence of certain patterns of team interaction that contribute to the understanding of team cognition, in general, and collaborative problem solving, specifically, through an integration of methods that measured team interaction dynamics and knowledge building as it occurred during a complex CPS task. This dissertation drew from recent theoretical and empirical work on Macrocognition in Teams and built on this by studying both verbal and non-verbal team interaction during complex CPS; a critical gap in this scientific area of inquiry. The present study also drew from work in cognitive science designed to study social and team interaction as a nonlinear dynamical system. In particular, a video frame-differencing method, an automated video analysis technique, was used to capture the bodily movements of participants and content coding of team communications was used to characterize their CPS processes. Time-lagged cross-correlation analysis was used to identify the degree of bodily-movement synchronization of teams. Recurrence quantification analysis, a dynamical systems method, was used to identify patterns of complementarity in teams’ communications.

The results of the present research showed that teams did synchronize their bodily movements and their communications could be characterized as complementary. However, there the results did not demonstrate a systematic relationship between these interaction measures and CPS performance on the Moonbase Alpha task. Video game experience, task knowledge, and cognitive load were significant covariates. While the current findings do not show a relationship between interaction-based measures and problem solving performance, this should not be taken as an indication that interaction, nor the methodological tools employed here, are not important for understanding collaborative problem solving. A number of limitations and future directions
have been identified to lay a foundation that advances research on team interaction dynamics during collaborative problem solving. Indeed, this research is but one small step in advancing the limited scientific knowledge in this crucial area.
APPENDIX A: LUNAR SURVIVAL MATERIALS
Lunar Survival Team Exercise

Process

- Give individuals slides 3, 4 and 5 of this packet.
  - Have individuals read the survival exercise and have each individual prioritize the items for survival in rank order (Step 1) (10-15 minutes)
- Once all individuals have done this and recorded their individual votes on their scoresheets, tell the group that they can discuss their answers together to learn from the collective wisdom in the room. Begin talking as a group to figure out how to reach a group decision about the ranking (25-30 minutes)
  - Group can decide whether to resolve differences through a collective vote or by further discussion
- On an overhead, record the group decisions and have each of the individuals fill this in on their scoresheet (Step 2 column)
- Supply the expert rankings from Side 5 and have group fill these in (Step 3)
- Have each participant compute the individual and group scores on the scoresheet (Steps 4 & 5)
- Record all the gains and losses and see how the collective wisdom of group improved the scores.
LUNAR SURVIVAL Exercise

In the following situation, you and your teammate’s “life” and “death” depends upon how well you can prioritize items for survival in a relatively unfamiliar environment. This problem is fictional, although the ranking to which you will compare your results was done by a number of space experts.

The Situation

You are a member of a lunar exploration crew originally scheduled to rendezvous with a mother ship on the lighted surface of the moon. Due to mechanical difficulties however, your ship was forced to land at a spot some 320 kilometers (200 miles) from the rendezvous point. During the re-entry and landing, much of the equipment aboard was damaged, and, since survival depends on reaching the mother ship, the most critical items available must be chosen for the 320 km trip.

Your Task

On the next page are listed the 15 items left intact and undamaged after landing. Your task is to rank these items according to their importance in aiding you to reach the mother ship, starting with “1” the more important, to “15” the least important. You and your teammate have agreed to stick together. Assume that all 15 items are in good condition.

Scoresheet Lunar Survival

<table>
<thead>
<tr>
<th>Items</th>
<th>Step 1 Indiv Ranking</th>
<th>Step 2 Team Ranking</th>
<th>Step 3 Expert Ranking</th>
<th>Step 4 Difference Ranking [1-3]</th>
<th>Step 5 Difference Ranking [2-3]</th>
<th>Your Score</th>
<th>Team Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compass</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First Aid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flares</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FM receiver</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Food concentrate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heat</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Map</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Matches</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milk</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxygen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parachute</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pistols</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raft</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rope</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total the absolute differences of Steps 4 and 5 ————————————> (the lower the score the better)

Your Score Team Score
## Lunar Survival Items

- Compass, magnetic
- First aid kit w/ hypodermic needles
- Flares, signal
- FM receiver/transmitter (solar-powered)
- Food concentrate
- Heating unit, portable
- Map (stellar map, moon’s constellations)
- Matches (1 box)
- Milk (1 case dehydrated milk)
- Oxygen (2 50kg tanks)
- Parachute silk
- Pistols (2 .45 caliber)
- Raft, Life (automatic inflating)
- Rope, Nylon (20 meters)
- Water (25 liters)

## Lunar Survival

### Ranking of Items by Experts

<table>
<thead>
<tr>
<th>Item</th>
<th>Rank</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oxygen</td>
<td>1</td>
<td>Fills respiration requirements</td>
</tr>
<tr>
<td>Water</td>
<td>2</td>
<td>Replenishes loss by sweating, etc</td>
</tr>
<tr>
<td>Map</td>
<td>3</td>
<td>One of principal means of finding directions</td>
</tr>
<tr>
<td>Food</td>
<td>4</td>
<td>Supply daily food required</td>
</tr>
<tr>
<td>FM receiver</td>
<td>5</td>
<td>Distress signal transmitter, possible communication with another ship</td>
</tr>
<tr>
<td>Rope</td>
<td>6</td>
<td>Useful in tying injured together, help in climbing</td>
</tr>
<tr>
<td>First aid kit</td>
<td>7</td>
<td>Oral pills or injection medicine available</td>
</tr>
<tr>
<td>Parachute</td>
<td>8</td>
<td>Shelter against sun’s rays</td>
</tr>
<tr>
<td>Raft</td>
<td>9</td>
<td>CO bottles for self propulsion across chasms, etc.</td>
</tr>
<tr>
<td>Flares</td>
<td>10</td>
<td>Distress call when line of sight possible</td>
</tr>
<tr>
<td>Pistols</td>
<td>11</td>
<td>Self propulsion devices could be made from them</td>
</tr>
<tr>
<td>Milk</td>
<td>12</td>
<td>Food mixed with water for drinking</td>
</tr>
<tr>
<td>Heating unit</td>
<td>13</td>
<td>Useful only if party landed on dark side</td>
</tr>
<tr>
<td>Compass</td>
<td>14</td>
<td>Probably no magnetized poles, therefore useless</td>
</tr>
<tr>
<td>Matches</td>
<td>15</td>
<td>Little or no use on moon</td>
</tr>
</tbody>
</table>
APPENDIX B: MOONBASE ALPHA TRAINING POWERPOINT
NASA Moonbase Alpha Training
Adapted from NASA Manual

Challenge Overview

- NASA has once again landed on the lunar surface and has established an outpost on the Moon called Moonbase Alpha. Utilizing solar energy and regolith processing, the Moon base has become self-sufficient and plans for further expansion are underway.

- In Moonbase Alpha, you have the role of an astronaut working to further human exploration and research. Returning from a research expedition, you witness a meteorite impact that severely damages the settlement’s life support capability. With only 25 minutes' worth of oxygen, you and your team must repair and replace equipment in order to restore the oxygen production to the settlement.

- Team coordination along with the proper use of your resources (player controlled robots, rovers, repair tools, etc.) are key to your success in restoring oxygen production. You will be scored on the time spent to complete the task, so you must communicate with your fellow astronaut in order to work effectively as a team.
Moonbase Alpha Settlement includes:

- **Solar Panel**: Absorbs energy from the sun.
- **Power Cables**: Carry energy to the power distributor.
- **Life Support System**: Uses power to create oxygen for living quarters.
- **Power Distributor**: Converts energy to usable power for life support system.
- **Living Quarters**: Houses research personnel.
- **Couplers**: Connect hoses.

---

**Equipment Overview**

- **Solar Panels**: Absorbs the Sun's energy and converts it to direct current electricity using a photovoltaic material (crystalline silicon).

- **Power Cables**: Transfer electricity from the solar panels to the power distributor and are linked together with couplers.

- **Power Distributor**: Using bridge circuits, converts the electric current from the solar panels into electrical power that can be used to create oxygen. Replacing the distributor components increases its efficiency of consuming energy, thereby providing more power to the life support system.
Equipment Overview

- **Solar Panels**: Absorb the Sun’s energy and convert it to direct current electricity using a semiconducting material (silicon).

- **Power Cables**: Transfer electricity from the solar panels to the power distributor and are linked together with connectors.

- **Power Distributor**: Using bridge circuits, converts the electric current from the solar panels into electrical power that can be used to create oxygen. Receiving the distributor components increases the efficiency of converting energy to power, thereby providing more power to the life support system.

Equipment Overview

- **Life Support System**: Uses electric power to heat the lunar regolith, breaking it down with its components. Once the breathable oxygen can be extracted and transferred to the living quarters. Receiving the system components increases the system’s efficiency, which increases the oxygen generation rate.

- **Rover**: Carries up to 2 people at a time, in addition to tools and equipment.
Equipment Overview

- **Equipment shed**: Contains tools, toolboxes, replacement equipment, and robots.

- **Command center**: Allows one player at a time to remotely control the rover and to view the settlement from various stationary cameras and each player's point-of-view.

Repair Tasks:

- While you are repairing damaged equipment, you will occasionally have the opportunity to bypass the structural damage and directly access the damaged circuitry or solar cells. If you repair the connections, you will reduce the total amount of time required for the repair.

- Once the repair task screen appears, trace the tip of the welding torch through the circuitry to solder in new connections. You have a limited amount of time to complete the connections, so work quickly and accurately.
  - **If you win**, you save a certain amount of repair time depending on the difficulty level.
  - **If you lose**, you will not save any repair time.
Solar Panels

- If damaged, the solar panel cannot absorb the sun’s energy
- To repair panels must be un-deployed (lowered) before the repairs can be made
- The welding tool will be required for repairs

Cables and Couplers

- If the coupler is damaged or if there is a section of cable that is not connected to a coupler, then the energy cannot be transferred beyond that point
  - A welding tool is needed for couplers
- After connecting a cable to a coupler, astronauts must use a wrench to secure the connection
  - If the cable is not secure, then there will be a flashing yellow light on the coupler
- Couplers that are getting power have a green light
Power distributor and life support system

- Cannot be accessed by people due to a hazardous coolant leak
- **Robots must be used to make repairs & replacements**: players need to get near the site and then deploy the robot, which they then control remotely
  - Both types of robots (welders & grabbers) will be needed
- **Fixing things in this area increases the efficiency of the conversion of energy to power (power distributor) and oxygen generation (life support system)**
- Components that need to be fixed or replaced include O2/N2 controllers, O2 generators, and CO2 filters

Rovers

- Can be driven by the person riding on the rover or remotely from the command center
- Powered by rechargeable solar batteries
How to Construct Robots:

- Robots can be permanently equipped with either an arm (allows for equipment pick-up / removal) or a welder (allows for equipment repairs).
- Drag the slider to choose the balance between speed and battery power. The speed of the robot determines its movement speed as well as how quickly it performs actions. (More speed means the robot moves and repairs faster, but it will also run out of battery much sooner.)
- Drop the case, select it and click “deploy” to control your robot.

Repairs

- Components can be repaired or replaced
- Damage levels
  - **Yellow**: light damage
  - **Orange**: moderate damage
  - **Red**: severe damage
  - **Black**: destroyed/not repairable
- The greater the damage, the longer it takes to repair
Schematic Map

The schematic map view displays energy flow, damage states for all components, and positions of players, robots and rovers.

Task Screen

1. **Player List**: View your teammate’s name and color.
2. **Loudness Item**: View the name and icon of the item that you are holding.
3. **Map**: View an overhead view of the lunar base. Pressing the 'B' key hides the map, the 'A' key minimizes the map to full screen, and the 'X' and 'Y' keys can be used to zoom the view in and out.
4. **Chat Box**: You will be communicating verbally with your team member during this task.
5. **Total**: Displays the total amount supplied to the living quarters. The game is won when this meter is completely full.
6. **Time**: Displays the time remaining.
7. **O2 Rate**: Displays the rate at which the O2 Total increases over time. Fruiting made in this base will increase your O2 rate.
Keyboard Controls

- **W**: Move forward.
- **S**: Move back.
- **A**: Move left (astronaut) / Turn left (rover & robot).
- **D**: Move right (astronaut) / Turn right (rover & robot).
- **Q**: Strafe left (robot only).
- **E**: Strafe right (robot only).
- **R / Num Lock**: Toggle autorun.
- **C**: Hide/Show chat box.
- **B**: Hide/Show map.
- **M**: Full screen map.
- `<`: Change map mode (normal/schematic).
- **+ Key**: Zoom in map.
- **- Key**: Zoom out map.
- **Enter**: Activate chat window.
- **Page Up**: Scroll chat box up.
- **Page Down**: Scroll chat box down.
- **Space Bar**: Jump (astronaut only).
- **Esc**: Open escape menu.
- **Arrow Key Up**: Move forward / Pan full screen map up.
- **Arrow Key Down**: Move back / Pan full screen map down.
- **Arrow Key Left**: Move left / Pan full screen map left.
- **Arrow Key Right**: Move right / Pan full screen map right.
- **7**: Settlement View - Focuses the map view on the lunar settlement.
- **9**: All View - Zooms the map view out to the fullest extent.

Controls
Keep in mind these tips:

- One tool or piece of equipment can be carried at a time; the heavier the object you're carrying, the slower you move.
- The toolbox allows you to carry more tools, but it is heavier.
- Robots with more speed move from place-to-place more quickly and make repairs faster, but could be harder to control. Battery power is important because several items will need to be fixed in the hazardous area.
- If there isn't at least one working energy pathway (solar array/cable) then no oxygen can be produced, so it may be best to fix one segment first so that some oxygen starts flowing.
- In most cases, it's quicker to replace red (critical) items than it is to repair them.
- Beacons can be used to identify areas to focus on (might save time so that each player isn't looking at the map all of the time) or meeting places (colors & names can be assigned).

To complete the task in under 25 minutes you need to communicate with your teammate on:

- How to divide work between team members
  - Decide which tools, replacement parts, and/or robots to take with you
  - Decide whether to repair or replace red items
- Setting Priorities:
  - Focus on repairing a complete energy flow path
  - Focus on increasing efficiency by repairing the power distributor and life support system components
  - Managing tradeoffs when creating a robot between speed and power
  - Updating your team member with regard to your status in completing a task, your location, or current items
  - Generating ideas for how you can complete your task in the least amount of time as possible
To quickest way to start restoring oxygen flow is to:

1. Repair a single solar panel
2. Ensure all power cables form a solid connection from the solar panel to the power distributor by connecting them to power couplers and tightening them with a wrench
3. Use the welding tool to repair all the connected power couplers
APPENDIX C: SOCIAL VALUE ORIENTATION ASSESSMENT
In this task we ask you to imagine that you have been randomly paired with another person, whom we will simply refer to as the “Other”. **This other person is someone you do not know and will not knowingly meet in the future.** Both you and the other person will be making choices by circling one of the letters: A, B or C. Your choices will produce points for both you and the other person. Likewise, the other person’s choice will produce points for them and for you. Every point has value: The more points you receive the better for you, and the more points the other person receives the better for them.

Here is an example of how this task works:

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>You get</strong></td>
<td>500</td>
<td>500</td>
<td>550</td>
</tr>
<tr>
<td><strong>Other gets</strong></td>
<td>100</td>
<td>500</td>
<td>300</td>
</tr>
</tbody>
</table>

In this example if you chose A you would receive 500 points and the other person would receive 100 points; if you chose B you would receive 500 points and the other person would receive 500 points; if you chose C you would receive 550 points and the other person would receive 300 points. So, you see that your choice influences both the number of points you receive and the number of points the other person receives.

Before you begin making choices, please keep in mind that there are no right or wrong answers – choose the option that you, for whatever reason, prefer the most. Also, remember that the points have value: the more you accumulate the better for you. Likewise, from the other’s point of view, the more points they accumulate the better for them.

Ask the experimenter now if you have any questions. Please turn over the page and complete the task.
For each of the nine choice situations below, circle A, B or C depending on the option you prefer the most.

<table>
<thead>
<tr>
<th></th>
<th>You get</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>480</td>
<td>540</td>
<td>480</td>
</tr>
<tr>
<td></td>
<td>Other gets</td>
<td>80</td>
<td>280</td>
<td>480</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>560</td>
<td>500</td>
<td>500</td>
</tr>
<tr>
<td></td>
<td>Other gets</td>
<td>300</td>
<td>500</td>
<td>100</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>520</td>
<td>520</td>
<td>580</td>
</tr>
<tr>
<td></td>
<td>Other gets</td>
<td>520</td>
<td>120</td>
<td>320</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>500</td>
<td>560</td>
<td>490</td>
</tr>
<tr>
<td></td>
<td>Other gets</td>
<td>100</td>
<td>300</td>
<td>490</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>560</td>
<td>500</td>
<td>490</td>
</tr>
<tr>
<td></td>
<td>Other gets</td>
<td>300</td>
<td>500</td>
<td>90</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>500</td>
<td>500</td>
<td>570</td>
</tr>
<tr>
<td></td>
<td>Other gets</td>
<td>500</td>
<td>100</td>
<td>300</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>510</td>
<td>560</td>
<td>510</td>
</tr>
<tr>
<td></td>
<td>Other gets</td>
<td>510</td>
<td>300</td>
<td>110</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>550</td>
<td>500</td>
<td>500</td>
</tr>
<tr>
<td></td>
<td>Other gets</td>
<td>300</td>
<td>100</td>
<td>500</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>480</td>
<td>490</td>
<td>540</td>
</tr>
<tr>
<td></td>
<td>Other gets</td>
<td>100</td>
<td>490</td>
<td>300</td>
</tr>
</tbody>
</table>
APPENDIX D: COGNITIVE LOAD RATING SCALE
In completing the previous Moonbase Alpha task, I invested:

1. Very, very low cognitive effort
2. Very low cognitive effort
3. Low cognitive effort
4. Rather low cognitive effort
5. Neither low nor high cognitive effort
6. Rather high cognitive effort
7. High cognitive effort
8. Very high cognitive effort
9. Very, very high cognitive effort

I experienced the previous Moonbase Alpha task as:

1. Not difficult at all
2. Very low difficulty
3. Low difficulty
4. Rather low difficulty
5. Neither low nor high difficulty
6. Difficult
7. Highly difficult
8. Very difficult
9. Very, very difficult

In completing the previous Moonbase Alpha task, I felt:

1. Very, very low frustration
2. Very low frustration
3. Low frustration
4. Rather low frustration
5. Neither low nor high frustration
6. Rather high frustration
7. High frustration
8. Very high frustration
9. Very, very high frustration
Select the correct answer to the following questions to the best of your ability. When you finish click next to view the correct answers.

1. What caused the damage to Moonbase Alpha?
   a. Solar flare
   b. Meteorite impact
   c. System malfunctions

2. Which of the following is NOT a component of the Moonbase Alpha settlement?
   a. Solar Panel
   b. Power Distributor
   c. Couplers
   d. Power Cables
   e. Life Support System
   f. Staging Base
   g. Living Quarters

3. Which of the following conveys the order of Moonbase Alpha’s energy distribution:
   a. Power Cables → Power Distributor → Solar Panels
   b. Power Distributor → Solar Panels → Power Cables
   c. Solar Panels → Power Cables → Power Distributor

4. Which of the following is NOT true about repairing damaged circuits (mini-game)?
   a. Trace the tip of the welding torch through the circuitry to solder in new connections.
   b. You have an unlimited amount of time to complete the repair tasks.
   c. If you win, you will save valuable repair time.
   d. You must first lower solar panels before you repair the circuitry.

5. Which of the following is NOT true of cables and couplers?
   a. If the cable is not secure, the coupler will have a yellow flashing light; use a wrench to secure the connection.
   b. A welding tool is needed to repair couplers.
   c. Couplers that are operating normally show a red light.

6. Why is teamwork essential to the successful completion of the mission?
   a. The task is too difficult to be completed without teamwork.
   b. The tasks in the mission are meant to be done with more than one player.
   c. Without dividing work between team members (deciding which tools, replacement parts, and/or robots to take with you), the mission will take too long to successfully complete.
   d. All of the above are correct.
7. Robots can be permanently equipped with either a(n) ___ or a(n) ____.
   a. Torch (for light), hammer (for repairs)
   b. Arm (equipment pick-up/removal), welder (repairs)
   c. Speed boost pack (for quicker movement), welder (for repair)

8. When a warning is displayed indicating that the area near the power distributor and the life support system is too dangerous for humans to repair broken equipment, which of the following tools must be used to repair those system components?
   a. Rover
   b. Equipment shed
   c. Welding tool
   d. Robot

9. It is often quicker to replace red items (severely damaged) than it is to repair.
   a. True
   b. False

10. When adjusting the speed and battery power of a robot, more speed means the robot will move and repair at a faster pace, but it will also run out of battery much sooner.
    a. True
    b. False
APPENDIX F: CORRELATION MATRICES
## Individual Level

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Age</td>
<td>19.24</td>
<td>2.16</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Gender</td>
<td>1.36</td>
<td>0.48</td>
<td>-0.108</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Race</td>
<td>1.78</td>
<td>1.26</td>
<td>0.093</td>
<td>0.074</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Video Game Familiarity</td>
<td>3.65</td>
<td>0.99</td>
<td>0.057</td>
<td>-0.619**</td>
<td>-0.081</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. 1st/3rd Person PC Game Familiarity</td>
<td>3.12</td>
<td>0.98</td>
<td>0.042</td>
<td>-0.440**</td>
<td>-0.084</td>
<td>0.639**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Moonbase Alpha Experience</td>
<td>1.99</td>
<td>0.11</td>
<td>-0.240*</td>
<td>0.081</td>
<td>-0.192*</td>
<td>-0.148</td>
<td>-0.211*</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Participant Familiarity</td>
<td>1.94</td>
<td>0.24</td>
<td>0.074</td>
<td>-0.020</td>
<td>0.036</td>
<td>0.063</td>
<td>-0.073</td>
<td>-0.027</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Duration of Participant</td>
<td>5.64</td>
<td>1.30</td>
<td>0.097</td>
<td>0.021</td>
<td>0.053</td>
<td>0.133</td>
<td>-0.020</td>
<td>-0.030</td>
<td>0.907**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Social Value Orientation</td>
<td>1.48</td>
<td>0.73</td>
<td>0.03</td>
<td>0.074</td>
<td>0.128</td>
<td>-0.092</td>
<td>-0.079</td>
<td>0.071</td>
<td>-0.110</td>
<td>-0.083</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Knowledge Assessment</td>
<td>9.16</td>
<td>0.97</td>
<td>-0.081</td>
<td>0.001</td>
<td>-0.086</td>
<td>0.182*</td>
<td>0.054</td>
<td>0.018</td>
<td>0.094</td>
<td>0.142</td>
<td>-0.078</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>11. Cognitive Load</td>
<td>5.55</td>
<td>1.44</td>
<td>0.023</td>
<td>0.125</td>
<td>-0.112</td>
<td>-0.148</td>
<td>-0.102</td>
<td>0.042</td>
<td>0.014</td>
<td>0.011</td>
<td>-0.008</td>
<td>-0.062</td>
<td>1</td>
</tr>
<tr>
<td>Team Level</td>
<td>Mean</td>
<td>SD</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>------</td>
<td>-----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Similarity</td>
<td>2.12</td>
<td>1.16</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Performance</td>
<td>29.41</td>
<td>31.82</td>
<td>0.16</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Synchrony</td>
<td>0.088</td>
<td>0.051</td>
<td>0.280</td>
<td>0.117</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Complementarity</td>
<td>13.01</td>
<td>1.85</td>
<td>0.148</td>
<td>0.132</td>
<td>0.116</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Cross-Level

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Similarity</td>
<td>2.12</td>
<td>1.16</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Performance</td>
<td>29.41</td>
<td>31.82</td>
<td>0.16</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Synchrony</td>
<td>0.088</td>
<td>0.05</td>
<td>0.280*</td>
<td>0.117</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Ave Video Game Fam</td>
<td>3.65</td>
<td>0.70</td>
<td>0.299*</td>
<td>.426**</td>
<td>.004</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Max Video Game Fam</td>
<td>4.19</td>
<td>0.73</td>
<td>0.142</td>
<td>.171</td>
<td>.151</td>
<td>.80**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Min Video Game Fam</td>
<td>3.12</td>
<td>0.93</td>
<td>.340*</td>
<td>.509**</td>
<td>-.113</td>
<td>.882**</td>
<td>.422**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Ave Knowledge Assessment</td>
<td>9.16</td>
<td>0.71</td>
<td>.250</td>
<td>.364**</td>
<td>.226</td>
<td>.116</td>
<td>.055</td>
<td>.132</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Maximum Knowledge Assessment</td>
<td>9.65</td>
<td>0.61</td>
<td>.126</td>
<td>.162</td>
<td>.232</td>
<td>.042</td>
<td>-.011</td>
<td>.073</td>
<td>.787**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Minimum Knowledge Assessment</td>
<td>8.67</td>
<td>1.02</td>
<td>.275*</td>
<td>.413**</td>
<td>.178</td>
<td>.137</td>
<td>.083</td>
<td>.141</td>
<td>.928**</td>
<td>.501**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Ave Cognitive Load</td>
<td>5.55</td>
<td>1.32</td>
<td>.212</td>
<td>-.480**</td>
<td>.009</td>
<td>-.175</td>
<td>-.166</td>
<td>-.134</td>
<td>-.186</td>
<td>-.080</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Max Cognitive Load</td>
<td>6.01</td>
<td>1.43</td>
<td>.176</td>
<td>-.458**</td>
<td>-.002</td>
<td>-.163</td>
<td>-.175</td>
<td>-.108</td>
<td>-.164</td>
<td>-.178</td>
<td>-.123</td>
<td>-.967**</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. MinCog Load</td>
<td>5.09</td>
<td>1.30</td>
<td>.236</td>
<td>-.467**</td>
<td>.021</td>
<td>-.175</td>
<td>-.142</td>
<td>-.153</td>
<td>-.097</td>
<td>-.181</td>
<td>-.027</td>
<td>.959**</td>
<td>.856**</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>13. Complementarity (%REC)</td>
<td>12.85</td>
<td>1.75</td>
<td>.148</td>
<td>.116</td>
<td>.100</td>
<td>.014</td>
<td>.140</td>
<td>-.014</td>
<td>.052</td>
<td>-.064</td>
<td>.111</td>
<td>.031</td>
<td>.048</td>
<td>.009</td>
<td>1</td>
</tr>
</tbody>
</table>

150
APPENDIX G: COLLABORATIVE PROBLEM SOLVING CODEBOOK
Collaborative Problem Solving Coding Scheme

**Code:**  
**IP—Information Provision**

**Brief Definition:** 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 (which are simple observations with no additional information/evaluation).

**Full Description:** IP statements always provide simple information. Simple information is 1) a fact that can be directly read from one place in the computer display or reference sheet, or 2) a ‘one bit’ statement of task information (e.g., equipment, a location, etc.). In these statements, there is no integration, analysis, or evaluation of the information in the actual utterance.

**When to use:** Use IP for any statements where someone is giving information that can be pulled from one place in the display. It does not matter if the person is reading from a display or recalling it (e.g., they remember someone else’s information or their own from a previous time), statements of simple information should be coded as IP.

Use IP codes for utterances when someone repeats information aloud (e.g., when talking to self) several times.

Use IP for ‘one bit’ statements of task information (e.g., equipment, a location, etc.). It does not matter if this ‘one bit’ statement requires complex analysis to provide, as long as there is no complex info in the statement.

Use IP when someone responds to a statement with the same information (i.e., an echo of an IP statement).

Use IP when people are providing information about the location of resources, equipment (e.g., ‘it’s right there’), or other players.

**When not to use:** Don’t use IP when the statement is complex in nature (that is, it integrates information from different sources) or it provides an evaluation of information (i.e., provides an opinion/evaluation of how good or bad the information is relative to the operation goals).

Don’t use IP when someone is providing simple information across a set of resources. For example, ‘all of these couplers are in bad condition’, or when summarizing an ability for a set of resources (e.g., ‘none of the tools over here can help’
Don’t use IP when someone is stating a goal. We are coding goal statements as GTO, not IP.

Don’t use IP when the utterance is a question.

Do not use IP to replace an original code if the statement is repeated by either participant.

Location is not exclusive to IP/IR if it is within an utterance involving further knowledge and/or an evaluation, which makes the information no longer simple.

### Examples:

<table>
<thead>
<tr>
<th>Positive Examples</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Oh, there’s a yellow light right next to you.”</td>
<td>This is information that can be pulled directly from one place in the display.</td>
</tr>
<tr>
<td>“Oh no, back is S.”</td>
<td>This is an example when someone was repeating information aloud while talking to themselves.</td>
</tr>
<tr>
<td>“M? Oh, okay.”</td>
<td>This is an example where someone is responding to the statement with the same information.</td>
</tr>
<tr>
<td>“I see you right there, but...”</td>
<td>This is a statement talking about the location of something.</td>
</tr>
<tr>
<td>“That’s a power distributor, on the right.”</td>
<td>This is an example where the statement is referring to information that can be accessed directly from the display and is also referring to the location of something.</td>
</tr>
<tr>
<td>“Red... Yeah.”</td>
<td>This is an example where the statement is a repetition of previously stated information.</td>
</tr>
</tbody>
</table>

### Negative Examples

<table>
<thead>
<tr>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>“How do we access a robot?”</td>
</tr>
</tbody>
</table>
“We need to find a wrench, right?” This statement is posed as a question of what the team needs to do, and therefore would be coded as GTO, not IP, even though the content is simple and could be interpreted as a repetition of information accessible from the screens.
**Code:**  
IR—Information Request

**Brief Definition:**  
Question utterances asking for a response of simple information about the task environment or situation, or questions asking for repetition of immediately preceding information.

**Full Description:**  
IR utterances always ask a question that requires simple information to answer. Simple information is a fact that can be directly read from one place in the information displays or reference sheets in the task. It does not require that the person sending or receiving the information to perform any type of integration, analysis, or evaluation of the information. It’s as if someone is asking someone else to perform a simple look up task. Additionally, IR utterances can be specific or general requests for clarification of immediately preceding information.

**When to use:**  
Use IR for any question utterances where someone is asking for simple information. It does not matter if someone responds with more complex information (or even if no one responds at all). You need to determine whether or not the response to the question can be read off one of the displays or requires more integration/evaluation.

Use IR for any specific and general questions asking for repetition or simple clarification of previous statements. This may depend upon the context of the utterance as well. General requests include things such as ‘Pardon?’, ‘What?’, ‘What was that?’, ‘Hm?’, etc.

Use IR when people are asking for information about the location of resources/equipment (e.g., couplers, solar panels, life support system) or other participants ("where are you?").

**When not to use:**  
Don’t use IR for statements, only questions.

Don’t use IR for questions that require a complex or evaluative response. These will likely be coded as KR—Knowledge Request.

Don’t use IR for questions about how to use the interfaces or displays. These will be coded as KR as they require knowledge that isn’t accessible from the displays themselves.

**Examples:**

<table>
<thead>
<tr>
<th><strong>Positive Examples</strong></th>
<th><strong>Rationale</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>“Where the one is completely destroyed?”</td>
<td>This is asking for the clarification of something that the other participant said.</td>
</tr>
</tbody>
</table>
“Is the tool shed over here?”

<table>
<thead>
<tr>
<th>Negative Examples</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>“How’d you do it?”</td>
<td>This is asking for knowledge that the other</td>
</tr>
<tr>
<td></td>
<td>participant has about their own actions and</td>
</tr>
<tr>
<td></td>
<td>thus would be coded as KR.</td>
</tr>
<tr>
<td>“So, how do we find out where...”</td>
<td>This is asking for knowledge about how to</td>
</tr>
<tr>
<td></td>
<td>operate the displays and thus would be</td>
</tr>
<tr>
<td></td>
<td>coded as KR.</td>
</tr>
</tbody>
</table>
**Code:**  
**KP—Knowledge Provision**

**Brief Definition:** Statements about the task environment or situation that provide either 1) an integration of more than one pieces of simple information, or 2) an evaluation or interpretation of the meaning, value, or significance of information with regard to the current subtask.

**Full Description:** KP statements are similar to IP statements; however, instead of providing simple information, they provide complex information. In contrast to simple information, *complex information* involves either 1) integrating information in a way such that the product of that integration is something not directly accessible from the information display (i.e., they combine information to create something new that can’t be read directly off of one of the computer displays), or 2) providing an evaluation of information in the displays relative to the team’s goals (i.e., they comment on the meaning or value of simple information).

**When to use:** Use KP for any statements where someone is providing complex information or an evaluation or opinion that is not directly related to a solution option previously generated by a team member.

Use KP statements for ‘anti-option’ statements—statements that describe what the team cannot do in a general sense.

Use KP (and KR) for utterances about the use of the computer interface. For example, “Okay, thing is, how do you turn around?” and “How do I drop it?” are KR statements. The responses to these questions are typically KP statements. On the other hand, “What are those things on the ground?” and “Yeah, I’m trying to repair it but it’s not letting me.” are not KR/KP statements because they involve simple information that is easily accessible from one screen. The first example would be coded as IR and the second example would be coded as SU/R because its purpose is to update the team on their current difficulties.

These statements are basically making sense of interface issues so will fall under the general rule of: if it’s about understanding the interface, it’s KP/KR. This includes information that could have been accessed directly from the training PowerPoint.

Use KP (and KR) for utterances about contingencies in information (e.g., if this is true, then it means X). These are basically discussing interpretations of meaning of information, but not specific options or evaluation of options.
Use KP for utterances listing what other tools/equipment are needed to complete an objective (e.g., I need a welding torch for it, too). Do not mistake with GTO (see “When not to use” for further clarification).

Use KP (and KR) for utterances where team members are discussing (amongst themselves) what they are and aren’t allowed to ask for from the experimenter.

Use KP if the team member is recalling information from the PowerPoint in regards to what tasks need to be completed (e.g., “It said we should repair the solar panels first.”)

**When not to use:** Don’t use KP for questions, only statements.

Don’t use KP for simple information statements.

Don’t use KP if someone is evaluating a solution that was generated.

Don’t use KP for statements that propose a specific action to be taken (This would be some form of Option Generation). KP statements will be declarative or evaluative in nature (i.e., they provide facts, knowledge, and evaluations).

**Examples:**

<table>
<thead>
<tr>
<th>Positive Examples</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>“You gotta right-click to drop it.”</td>
<td>This is a statement answering a question about a problem with the interface.</td>
</tr>
<tr>
<td>“Oh, this is how we talk.”</td>
<td>This is a comment about the use of the interface.</td>
</tr>
<tr>
<td>“You need both hands free.”</td>
<td>This is a comment about the use of the interface.</td>
</tr>
<tr>
<td>“To look around you use the mouse.”</td>
<td>This is a comment about the use of the interface.</td>
</tr>
<tr>
<td>“Yeah, ‘cause the stuff is red, so when it’s red you have to replace it.”</td>
<td>This is an example of the communication of information that could have been directly accessed from the training PowerPoint.</td>
</tr>
<tr>
<td>“I need a welder.”</td>
<td>This is an utterance where the participant is listing what other tools/equipment is needed to complete an objective.</td>
</tr>
<tr>
<td><strong>Negative Examples</strong></td>
<td><strong>Rationale</strong></td>
</tr>
<tr>
<td>----------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>“Yeah, so we can replace those.”</td>
<td>This statement is proposing a specific course of action, with only a subset of the solution components, so it would be coded as OG-P.</td>
</tr>
</tbody>
</table>
Code: **KR-Knowledge Request**

**Brief Definition:** Question utterances that request a complex information response about the task environment or situation: to answer the question, the response should provide either 1) an integration of more than one piece of simple information, or 2) an evaluation or interpretation of the meaning, value, or significance of information within the current subtask.

**Full Description:** KR utterances always ask a question that requires a complex information response about the task environment or situation. In contrast to simple information, complex information involves either 1) integrating information in a way such that the product of that integration is something not directly accessible from the information display (i.e., they combine information to create something new that can’t be read directly off the computer display), or 2) providing an evaluation of information in the displays relative to the team’s goals (i.e., they comment on the meaning or value of simple information). It does not matter if someone responds with simple information (or no one responds at all). You have to determine whether answering the question requires integration or evaluation of information or not.

**When to use:** Use KR for questions requiring complex information responses (integration, evaluation, analysis).

Use KR for utterances questioning what other tools/equipment are needed to complete an objective (e.g., what else do we need?).

Use KR for general requests for resource information (e.g., ‘do you have anything around the solar panels?’)

Use KR for statements that are about making sense of the computer interface (e.g., figuring out how to click on objects or what buttons are needed to perform a certain action).

Use KP (and KR) for utterances about the use of interfaces. For example, “How do I? How do you work on...?” and “How do I drop it?” are KR statements. The responses to these questions are typically KP statements. On the other hand, “What are those things on the ground?” and “Yeah, I’m trying to repair it but it’s not letting me.” are not KP/KR statements because they involve simple information that is easily accessible from one screen. The first example would be coded as IR and the second example would be coded as SU/R because its purpose is to update the team on their current difficulties.
Use KR (and KP) for discussing what the limits of the game are. For example, “Can we fly?” are KR statements. The responses to these questions are typically KP statements.

Use KR (and KP) for utterances where team members are discussing (amongst themselves) what they are and aren’t allowed to ask for from the experimenter.

**When not to use:**

Don’t use KR for statements, only questions.

Don’t use KR for questions requiring simple information responses.

### Examples:

<table>
<thead>
<tr>
<th>Positive Examples</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Where’d you get a new coupler from?”</td>
<td>This is a question that requires integration of simple information to answer.</td>
</tr>
<tr>
<td>“How do I pick up my wrench?”</td>
<td>This is a question asking about a problem with the interface.</td>
</tr>
<tr>
<td>“Okay, well, what’s red right now? Do you know what’s red?”</td>
<td>This question might be coded as IR, because it’s asking about information that can be drawn from the display; however, the addition of the second question, where one participant asks the other if they know what’s red makes it evaluative in nature and therefore it should be coded as KR.</td>
</tr>
<tr>
<td>“So, how do we find out where the solar panels are?”</td>
<td>This is a question about the use of the interface.</td>
</tr>
<tr>
<td>“How do you replace it?”</td>
<td>This is a question about the use of the interface.</td>
</tr>
<tr>
<td>“Yeah. How do you pick this up? I need to pick it back up.”</td>
<td>This is a question about the use of the interface. The combination of the first half of this statement with the second might appear confusing, however the context of the second half, which otherwise would be coded as GTO (self-directive, a comment on what needs to happen), is as a continuation</td>
</tr>
</tbody>
</table>
of the original question which makes it KR.

<table>
<thead>
<tr>
<th><strong>Negative Examples</strong></th>
<th><strong>Rationale</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>“What are those on the ground? Are those welders?”</td>
<td>This is a question requiring a simple information answer.</td>
</tr>
</tbody>
</table>
**Code:**  
**OG-F—Option Generation - Full**

**Brief Definition:** Statements explicitly proposing a complete or near complete solution option for the team to consider and evaluate—a sequence of actions intended to accomplish part of the task. A complete solution includes reference to specific actions, tools, system components, and actors.

**Full Description:** OG-F utterances **propose** a complete or near complete solution for the team to consider executing. A solution is a set sequence of actions intended to meet one of the operation objectives. A complete solution includes four components: 1) **what** to do (what action are they proposing to be completed?), 2) **which component** (which components are they proposing the actions to be performed upon?), 3) **why** (for what purpose are they making the proposal? what component needs action? E.g., is the coupler destroyed? Is the solar broken? ) and 4) **how** (which tool are they proposing for use?). To be coded as OG-F, the utterance should include all of these components and it **should be suggestive not directive**.

OG-F statements are generally action statements that involve generating/proposing a potential solution (e.g., ‘I can take the wrench to the solar panel on the right so that I can repair it.’). OG-F statements can be stated as questions. For example, ‘Why don’t you take the welding torch over by the tubes, and then I’ll take the rest of the tools on the rover so that we can repair the couplers?’ is proposing an option to the group. The key for OG-F statements is that they 1) propose a potential solution and, 2) involve a complete sequence of actions to meet a specific goal (e.g., fixing a coupler). Further, they should be things that are proposed for the team to consider rather than directions for the team to execute.

**When to use:** Use OG-F for any sequence of actions for addressing one of the operation objectives and containing all of the option components (i.e., what, which, why, and how) specified explicitly.

Use OG-F when the utterance is in reference to the team level and what they should consider doing next (as opposed to just telling a specific teammate what they should do). These will be suggestive and not directive.

Use OG-F when an utterance proposes that one team member perform a given action and that they will do another action, why, and how (e.g., Why don’t you attach the hoses with the wrench and I will fix the solar panels with the welding tool, so we can get the oxygen flowing?”

**When not to use:** Don’t use OG-F to code statements implying a sequence of actions where not all of the solution components are specified explicitly (i.e., what, where, and why but not how).
Don't use OG-F when it should be GTO.

<table>
<thead>
<tr>
<th>Examples:</th>
<th><strong>Positive Examples</strong></th>
<th><strong>Rationale</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>“One of us should get a replacement coupler and replace that red coupler over there so that we can get some oxygen going.”</td>
<td>This is a suggestive statement that includes what (replace the red coupler), why (can get some oxygen going), and how (get a replacement coupler).</td>
<td></td>
</tr>
<tr>
<td>“If you could bring me a wrench, because I’m gonna need that to disconnect this coupler near the solar panel since it is black and needs to be replaced.”</td>
<td>This involves a complete solution: what (disconnect a coupler), where (the hose from the coupler), why (the coupler is black), and how (by using a wrench).</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Negative Examples</strong></th>
<th><strong>Rationale</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>“If you could bring me a wrench, because I’m gonna need that to disconnect the coupler.”</td>
<td>This does not encompass the 4 components listed above, so this statement should be labeled as OG-P.</td>
</tr>
</tbody>
</table>
Code:  

**OG-P—Option Generation - Part**

**Brief Definition:**  Statements that provide an incomplete solution option for the team to consider and evaluate — 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.

**Full Description:**  OG-P utterances propose an incomplete solution. A solution is a set sequence of actions intended to meet one of the operation objectives. A complete solution includes four components: 1) what to do (what action are they proposing to be completed?), 2) which component (which components are they proposing the actions to be performed upon?), 3) why (for what purpose are they making the proposal? what component needs action? E.g., is the coupler destroyed? Is the solar broken?) and 4) how (which tool are they proposing for use?). To be coded as OG-F, the utterance should include all of these components and it should be **suggestive not directive**. Further, they should be things that are proposed for the team to consider rather than directions for the team to execute.

**When to use:**  Use OG-P when the utterance proposes two or less of the four solution components (e.g., action, location, reason, and tool) to meet an operation objective.

Use OG-P when the utterance is in reference to the team level and what they should consider doing next. These will be suggestive and not directive.

Use OG-P when an utterance proposes that one team member perform a given action and that they will do another action, for a particular reason (e.g., Why don’t you attach the hoses and I will fix the solar panels, so we can get the oxygen flowing?”

**When not to use:**  Don’t use OG-P statements for utterances with all four solution components.

**Examples:**

<table>
<thead>
<tr>
<th>Positive Examples</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Yeah, so we can replace those.”</td>
<td>This is solution proposal that includes only two components, action and location. In this case the action is replacing the broken objects and the location is the ‘those’ which</td>
</tr>
</tbody>
</table>
refers to a broken component of the Moonbase.

“Yeah, um, if you could bring me a wrench, because I’m gonna need that to disconnect the coupler” (T5-92)

<table>
<thead>
<tr>
<th>Negative Examples</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Should I get tools, then, or should I get replacement?”</td>
<td>This player is asking what he should be working on, so it should be coded GTO.</td>
</tr>
<tr>
<td>Code:</td>
<td><strong>SEval—Solution Evaluation</strong></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td><strong>Brief Definition:</strong></td>
<td>Utterances that 1) compare different potential solutions, 2) provide support or criticism of a single potential solution, or 3) ask for an evaluation of a potential solution.</td>
</tr>
<tr>
<td><strong>Full Description:</strong></td>
<td>Solution Evaluation (SEval) utterances support, criticize, or ask for an evaluation of an option. Support and criticism can be specific (e.g. “If you fix that one, the pipes should be good”) or general (e.g. “I don’t know if that’s even worth it”) and can involve the direct comparison of different options or refer to a single potential solution.</td>
</tr>
<tr>
<td><strong>When to use:</strong></td>
<td>Use SEval when an utterance refers to the pros or cons of a solution option. Use SEval when people are comparing two different solution options in terms of quality (i.e., speed, ease of executing). Use SEval for utterances giving a final confirmation of a solution option (e.g. context specific utterances where an OG-P preceded a statement like “yes, let’s do that” in which case it is more than simple agreement, but is confirmation of the previously generated solution) Use SEval for utterances where there is an option and an evaluation in the same utterance. Use SEval for utterances where indifference, or lack of a preference toward an option or multiple options is expressed. For example, “I guess” and “I don’t care” for conditions where the prior utterance has generated an option.</td>
</tr>
<tr>
<td><strong>When not to use:</strong></td>
<td>Don’t use SEval to code statements where people are proposing, modifying, or clarifying options. Don’t use SEval for utterances that provide simple agreement or disagreement (i.e., S statements). SEval utterances provide more than just ‘yes or no’ type responses.</td>
</tr>
<tr>
<td><strong>Examples:</strong></td>
<td><strong>Positive Examples</strong></td>
</tr>
<tr>
<td></td>
<td>“It seems easier”</td>
</tr>
</tbody>
</table>
"Ummm, it doesn't matter either way, like..." (T5-92)  Statement is expressing indifference toward possible solutions.

“That way you won’t have to go back and forth?”  Evaluating the benefits of a single solution

“It’s probably going to slow me down a little bit, but it’s gonna have everything we need in it”  Weighs pros and cons of a solution

<table>
<thead>
<tr>
<th>Negative Examples</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Nah.”</td>
<td>This statement is a simple disagreement.</td>
</tr>
<tr>
<td>“Yeah, so we can replace those.”</td>
<td>This statement is proposing a solution option.</td>
</tr>
<tr>
<td>“Well that’s because mine was light damage”</td>
<td>Not SEval, because it is reflecting on past events. Should be coded R.</td>
</tr>
</tbody>
</table>
**Code:**  
*GTO—Goal and Task Orientation*

**Brief Definition:** 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 that an individual team member need to take to address a goal. These statements direct what the he or she should do next or later in the future. This includes self-references for an individual.

**Full Description:** GTO utterances are about high-level goals—things the team members need to do—and things the team members need to do to reach these goals (i.e., tasks). Consequently, these are future-oriented statements. GTO utterances include both providing and questioning the goals of the team.

**When to use:** Use GTO for statements where the person is telling a team member to focus attention on completing a task.

Comments on what needs to happen will be GTO statements unless they are in reference to the tools and resources needed which are KP.

Use GTO for assertive or command statements (e.g., “Ok, can you find out X.”)

Use GTO for utterances in which one participant asks the other if they would or if they want to work on a particular task. (e.g. “Do you want to work on the couplers?” or “Do you want me to go to the shed?”)

Use GTO for self-directing statements (e.g., “I’ll do X”)

Use GTO for utterances commenting on how to do a specific task in terms of the steps necessary to achieve it. Use KP for utterances commenting on the technical procedure for completing a specific task (e.g., which way to move a mouse or how to use a robot). The utterance can be characterized as to whether the statements are propositional/suggestive (OG) or directional/assertive (GTO). OG statements are put forth for the team members to consider/evaluate. GTO statements are put forth to direct the team towards completing their task/accomplishing their goal.

Use GTO for utterances where people ask what they should be doing.

Use GTO for utterances where a team member gives sequential commands. For example, “First we need do X, then X.”

Use GTO for utterances where a team member is ‘indirectly’ guiding another team member to do some task (e.g., ‘so someone needs to repair that coupler’)
Use GTO if the team member is including information that is goal/task oriented in regards to what tasks need to be completed.

Use GTO for utterances that may include any of the components of a solution (referred to in OG-F or OG-P), but GTO statements are directed and not characterized by being an option that the other team member consider/evaluates.

**When not to use:** Don’t use GTO for statements when someone is referring to what is happening right now or what they are doing right now. These are likely SU/R statements.

Don’t use GTO for statements proposing a solution option. GTO statements are about actions team members have to perform, and not about a proposed plan option for the team to consider. It can be difficult to distinguish between GTO and OG statements, but an **OG statement is a proposed sequence of actions for meeting an objective** and **GTO statements are more general**—they mention something that has to be done to execute an option. Further GTO statements should primarily be at an individual level (e.g., you do this) and OG statements should be at the team level (e.g., how about we do this because...OR why don’t you do this and I will do this because...)

**Examples:**

<table>
<thead>
<tr>
<th><strong>Positive Examples</strong></th>
<th><strong>Rationale</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>“And I’m gonna go pick up my wrench.”</td>
<td>This is a future oriented self-directive.</td>
</tr>
<tr>
<td>“You want me to just drop it?”</td>
<td>This statement mentions something that has to be done to execute an option.</td>
</tr>
<tr>
<td>“All right, let’s go to the black one.”</td>
<td>This statement is suggesting the group focus attention on completing a task.</td>
</tr>
<tr>
<td>“I need to go there again.”</td>
<td>This is a comment on what needs to happen.</td>
</tr>
</tbody>
</table>
“I need, um, welders right? So I need to go back.” This statement could be coded as either KP or GTO. The first half of the sentence is asking what additional tools are needed to complete the task, but in this context it serves as a reason for why the participant needs to go back to retrieve the other tool so it should be coded as GTO.

“So you want me to get one too?” This could fall into two aspects of GTO. The first being that one participant is asking the other if they would like them to work on a particular task, and the second being that the participant is asking what they should be working on.

“I’m just gonna focus on this, so…” This is a self-directive statement letting the other participant know what they’ll be working on. It’s future-oriented so it wouldn’t be coded as SU.

“I guess I’ll go over to the power cables.” This is a self-directive statement about what the participant will be doing.

“Okay, let me drop my equipment.” This is a comment on the steps necessary to complete a certain task.

“So you have to drop it and pick it up.” This is a comment on the steps necessary to complete a certain task.

**Negative Examples**

<table>
<thead>
<tr>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>This is a statement where a component of the solution option is being generated and thus would be coded OG-P.</td>
</tr>
<tr>
<td>This is a specific sequence of actions required in order to meet the objective, where GTO is more a general statement of goals.</td>
</tr>
</tbody>
</table>
Code: **SU—Situation Update**

**Brief Definition:** Statements that provide updates about what has been completed/what the team member is currently working on.

**Full Description:** SU statements comment on what is presently occurring with the team and the task. This includes self-references for an individual (e.g., what that person is currently doing), references to task completion (e.g., “hose is fully extended” as well as issues they are addressing (e.g., “it doesn’t let me bring the other one over”). By definition, SU statements should primarily be **present tense, unless they mention something they just completed (e.g., “I just repaired a coupler”).**

**When to use:**

- Use SU when team members are talking about themselves as a whole or as individuals and discussing what is happening right now.
- Use SU when team members are talking about what is happening with their task (e.g., success or failure with the mini games, effective/ineffective use of mouse or keyboard).
- Use SU when team members are talking about the status of executing their plan (e.g., what they currently have completed or are working on completing).
- Use SU when a team member is updating another team member on tasks they’ve completed.
- Use SU for discussions of time limits, remaining time, and oxygen levels in the operation.
- Use SU for utterances that are listing resources at a location at the end of an operation. That is, some of the statements that would normally be considered ‘one bit’ information statements (e.g., ‘2 couplers) can be SU/R when they are providing the team an update/verification of what has been repaired.

**When not to use:**

- Don’t use SU statements that comment on what needs to happen. Comments on what needs to happen will be GTO statements.
- Don’t use SU for utterances critiquing or evaluating the team’s past performance (these are R), only commenting on task completion/status.

**Examples:**

<table>
<thead>
<tr>
<th>Positive Examples</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>“I’m working on the couplers.”</td>
<td>States what player is currently doing</td>
</tr>
</tbody>
</table>
“Oh no, we died.” Statement provides an update of the current progress towards the main objective.

“I have the wrench” Statement is discussing the status of their plan and is not referencing what tool is needed to complete the task (which would be KP).

“I’ve fixed three, I have a whole row working” Statement addressing completion of tasks

<table>
<thead>
<tr>
<th>Negative Examples</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>“I need to go there again.”</td>
<td>This is a comment on what needs to happen and will be coded as GTO.</td>
</tr>
<tr>
<td>“All right, this is a lot more complicated than I thought it would be.”</td>
<td>This is a critique/evaluation of past work (the statement that it’s more complicated than they thought implies that they’re not preceding at the preferred pace) so it would be coded as R.</td>
</tr>
</tbody>
</table>
**Code:**  
**SR—Situation Update Request**

**Brief Definition:** Statements that ask about what the team is currently doing or what is currently happening with the simulation.

**Full Description:** SR statements ask about what is presently occurring with the team and the task. SR statements should be primarily warrant a present tense, or they could be past tense if they are asking about a situation that the team member just experienced.

**When to use:** Use R when team members are asking for updates about the progress of a task.

**When not to use:** Don’t use SR when team members are stating their progress or the progress of their tasks in general, as these will be coded as SU.

**Examples:**

<table>
<thead>
<tr>
<th>Positive Examples</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Did you get ahold of a robot?”</td>
<td>This is an example where someone is asking for more information about what is presently occurring with the team and the task.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Negative Examples</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>“I’ve fixed three, I have a whole row working.”</td>
<td>This is a situation update/update of the progress of the participant towards a goal.</td>
</tr>
<tr>
<td>“Do you want me to bring the tools to you?”</td>
<td>This is a GTO statement because it is one person asking the other what they should be working on.</td>
</tr>
</tbody>
</table>
**Code:**  
*R—Reflection*

**Brief Definition:** Utterances that provide or ask for a critique or evaluation of the performance of the team as a whole or of individual members.

**Full Description:** R utterances comment on or question the quality of the group’s performance or propose alternative ways of doing things that in hindsight might have been more effective. R is a reflection of things the team explicitly chose to do and how that has affected their performance or will affect their performance.

**When to use:** Use R for utterances that comment on the quality of work that has been accomplished, or discuss how the team has been working together (i.e., its processes). Reflection, by definition implies that the team is referring to something they have already done, so should be *past tense*.

*Use R for utterances that provide a projected outcome, i.e. using the status of current performance to reflect on future outcomes.*

*Use R for utterances that comment on what the team should have done or potentially could have done differently.*

*Use R if a team member referred to the quality of the task completion. After that, it might need to be broken down further (i.e., I just completed the soldering minigame (SU), and I did it very poorly (R)*

**When not to use:** Don’t use R for utterances communicating what the team is currently doing. These are likely SU/R statements.

Don’t use R for utterances that just state what tasks have been completed but do not provide evaluation of the quality of that work.

Don’t use R for utterances where people are using ‘I thought...’ utterances to communicate their understanding (or lack of understanding) about the situation. For example, ‘I thought it was to the right’ is an IP statement and “I thought they were operational” is a KP statement.

**Examples:**

<table>
<thead>
<tr>
<th>Positive Examples</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>“I feel like we might fail”</td>
<td>This is a reflective statement because it’s commenting on the quality of the work that has been accomplished and it is projecting an outcome (they might fail).</td>
</tr>
<tr>
<td>“Yeah, I should’ve probably gotten the wrench too.”</td>
<td>This is a comment on what the team (in this case a team member) should have done or could have done differently.</td>
</tr>
<tr>
<td><strong>Negative Examples</strong></td>
<td><strong>Rationale</strong></td>
</tr>
<tr>
<td>-----------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>“The shed says “back”, so, that’s where I’m trying to go.”</td>
<td>This is communicating what the team is currently doing, so it would be coded as SU.</td>
</tr>
<tr>
<td>Code:</td>
<td><strong>S</strong>— <em>Simple agreement/disagreement and Acknowledgement</em></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td><strong>Brief Definition:</strong></td>
<td>Simple agreement/disagreement utterances are expressions of agreement or disagreement with no rationale provided. Acknowledgements are utterances providing recognition of receipt of communication.</td>
</tr>
<tr>
<td><strong>Full Description:</strong></td>
<td>Simple agreement/disagreement utterances provide the equivalent of ‘yes/no’ responses to questions or statements. Acknowledgments are similar in that they are brief responses to statements or questions, but do not include further elaboration or meaning beyond simply responding. Look at the context of the statement so as to decide if S should be used or not.</td>
</tr>
<tr>
<td><strong>When to use:</strong></td>
<td>Use S for any simple yes or no responses or an equivalent. Use S for acknowledgement phrasings such as ‘let me see’, ‘ok’, ‘wait’, etc. Use S when team members echo one other to acknowledge that person and indicate an understanding of what they said.</td>
</tr>
<tr>
<td><strong>When not to use:</strong></td>
<td>Don’t use S for any utterance that includes an acknowledgement followed by substantive content such as ‘Yeah, the couplers are over there.’</td>
</tr>
<tr>
<td><strong>Examples:</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Positive Examples</strong></td>
</tr>
<tr>
<td></td>
<td>“Yeah!”</td>
</tr>
<tr>
<td></td>
<td>“Yeah, I think.”</td>
</tr>
<tr>
<td></td>
<td><strong>Negative Examples</strong></td>
</tr>
<tr>
<td></td>
<td>“Yeah, that’s what I need.”</td>
</tr>
</tbody>
</table>
Code: INC/F/EX—Incomplete/Filler/Exclamations

Brief Definition: 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. Incomplete utterances are statements that have no explicit meaning because they are missing one or more critical components of grammar: subjects, verbs, or objects.

Full Description: Incomplete utterances are usually false starts to communication that do not have any real meaning. These are not to be confused with ‘one bit’ statements coded as IP. These are not grammatically correct or necessarily a complete thought, but they are task related information. Incomplete utterances occur most frequently when someone begins speaking but does not finish the thought resulting in a statement with no meaning.

Fillers, or hedges, are place holders in communication. They fill gaps in between substantive speech. Examples include: "uh", "er" and "um". Additionally, words or phrases that can be substantive at times can also be used as fillers. For example, “Ok”, “Let me see”, and “Wait a minute” can all be filler statements or substantive communication in different contexts. It is up to you as a coder to determine if this is a ‘place holder’ or if it is an effort to communicate actual information. Usually, if these statements are in response to another utterance, they are substantive and would be coded as S.

Exclamations are single word or short phrase interjections used to communicate an emotional reaction to an event or a general situation. They have no meaning outside of communicating emotional content.

When to use: Use INC/F/EX for any utterances where the person is using a few words to express an emotional state or reaction (usually frustration or anger, but also excitement or joy) and no other explicit meaning.

Use INC/F/EX for utterances that end in negations of the entire utterance (e.g., never mind, forget it, etc.).

Use INC/F/EX for utterances that the transcriber can’t make out or are inaudible.

Use INC/FEX for laughter.
When not to use: Don’t use INC/F/EX when the meaning of the utterance is conveyed even if they don’t complete their whole statement. Don’t use INC/F/EX for any statements where there is a false start or trailing end attached to a complete thought. If any part of the utterance is complete and has meaning, code that meaning and ignore the incomplete aspects. Don’t use INC/F/EX if you think an utterance is an exclamation, but it has explicit meaning outside of expressing emotions. If it has explicit meaning outside of expressing an emotional reaction/state it is more than an exclamation. Don’t use INC/F/EX for ‘one bit’ IP utterances—utterances sharing or repeating task related information.

Examples:

Positive Examples | Rationale
--- | ---
“Ahh!” | Exclamations are single word or short phrase interjections used to communicate an emotional reaction to an event or a general situation. They have no meaning outside of communicating emotional content.
“Maybe there’s a way we can...” | This statement begins the expression of a potential solution but the cut off the inclusion of any of the components makes it an incomplete statement instead.
“There we go” | This is a filler statement.
“Okay” | This can be coded as INC/F/EX if it is not used as a simple agreement or confirmation of receipt of information.
<Laughter> | Laughter is expresses a specific emotion, therefore it should be coded as INC/F/EX.
“Alright” | This can be coded as INC/F/EX if it is used as a filler statement or simple.
<table>
<thead>
<tr>
<th><em>Negative Examples</em></th>
<th><em>Rationale</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>“I don't understand, how do I .... Grrrr!”</td>
<td>Although the original sentence trails off and leads into an exclamation, the player is clearly expressing uncertainty, so the code that should be used is UNC.</td>
</tr>
<tr>
<td>“Ummm, it doesn't matter either way, like...”</td>
<td>Sentence trails off, but the beginning of the sentences expresses indifference toward possible solutions, so the code that should be used is SEval.</td>
</tr>
</tbody>
</table>
**Code:**  
*T/OT—Tangent/Off-task*

**Brief Definition:**  
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.

**Full Description:**  
Tangent and Off-task utterances are those that deal with anything not directly related to task performance. This includes talking about things outside of the experiment, commenting on the experiment itself (e.g., what the participant’s think the experiment is about or ‘what we’re doing to them’), or jokes and sarcasm about aspects of the task.

**When to use:**  
Use this code for utterances that deal with anything not directly related to task performance.

**When not to use:**  
Don’t use this code when statements have to do with the task at hand.

**Examples:**

<table>
<thead>
<tr>
<th>Positive Examples</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>“How do people finish this in 25 minutes?”</td>
<td>This is not directly related to task performance.</td>
</tr>
<tr>
<td>“How does anybody get anything done?”</td>
<td>This is not directly related to task performance.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Negative Examples</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;Laughter&gt;</td>
<td>This is an expression of a specific emotion and would be coded INC/F/EX</td>
</tr>
</tbody>
</table>
Approval of Human Research

From: UCF Institutional Review Board #1
FWA0000351, IRB0000115

To: Travis J. Wiltshire and Co-PIs: Florian G. Jessch & Stephen M. Fiore

Date: July 29, 2014

Dear Researchers,

On 07/29/2014, the IRB approved the following human participant research until 7/28/2015 inclusive:

Type of Review: UCF Initial Review Submission Form
Project Title: Understanding Interaction Dynamics during Collaborative Problem Solving
Investigator: Travis J. Wiltshire
IRB Number: SBE-14-10439
Funding Agency: N/A
Grant Title: N/A
Research ID: N/A

The scientific merit of the research was considered during the IRB review. The Continuing Review Application must be submitted 30 days prior to the expiration date for studies that were previously expedited, and 60 days prior to the expiration date for research that was previously reviewed at a convened meeting. Do not make changes to the study (i.e., protocol, methodology, consent form, personnel, site, etc.) before obtaining IRB approval. A Modification Form cannot be used to extend the approval period of a study. All forms may be completed and submitted online at http://iris.research.ucf.edu

If continuing review approval is not granted before the expiration date of 7/28/2015, approval of the research expires on that date. When you have completed your research, please submit a Study Closure report in IRIS so that IRB records will be accurate.

Use of the approved, stamped consent document(s) is required. The new form supersedes all previous versions, which are now invalid for further use. Only approved investigators (or other approved key study personnel) may solicit consent for research participation. Participants or their representatives must receive a copy of the consent form(s).

All data, including signed consent forms if applicable, must be retained and secured per protocol for a minimum of five years (or HIPAA applies) past the completion of the research. Any links to the identification of participants should be maintained and secured per protocol. Additional requirements may be imposed by your finding agency, your department, or other entities. Access to data is limited to authorized individuals listed as key study personnel.

In the conduct of this research, you are responsible to follow the requirements of the Investigator Manual.

On behalf of Sophia Dzegilewski, Ph.D., L.C.S.W., UCF IRB Chair, this letter is signed by:

[Signature]

IRB Coordinator
REFERENCES


Bates, D.M.,Maechler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using S4 classes. R package version 0.999999-0.


teams. In Digital Ecosystems and Technologies (DEST), 2013 7th IEEE International Conference on (pp. 43-48). IEEE.


