FURTHERING HUMAN-ROBOT TEAMING, INTERACTION, AND METRICS THROUGH COMPUTATIONAL METHODS AND ANALYSIS

A Dissertation
Presented to
The Academic Faculty

By

Lanssie Mingyue Ma

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Computational Science and Engineering & Aerospace Engineering

Georgia Institute of Technology

May 2019

FURTHERING HUMAN-ROBOT TEAMING, INTERACTION, AND METRICS THROUGH COMPUTATIONAL METHODS AND ANALYSIS

Approved by:

Dr. Karen M. Feigh, Advisor
School of Aerospace Engineering
Georgia Institute of Technology

Dr. Richard Fujimoto
School of Computational Science and Engineering
Georgia Institute of Technology

Dr. Ashok Goel
School of Computer Science
Georgia Institute of Technology

Dr. Sonia Chernova
School of Computer Science
Georgia Institute of Technology

Dr. Terry Fong
Director of Intelligent Robotics
NASA Ames

Date Approved: 28 February, 2019

Shoot for the moon. Even if you miss, you’ll land among the stars.

Norman Vincent Peale

For the pioneers of research who make today's work possible and for the future pioneers who have to make sense of it.

ACKNOWLEDGEMENTS

Firstly, I would like to thank my advisor, Dr. Feigh for always being bold and open minded to my wild ideas. Her continuous support and flexibility with me as I jumped from ideas to internships was paramount to the discoveries I have made and people I have met today. Her bar for quality has made me a better researcher and challenges me to raise my own. I sincerely appreciate the endless rounds of polishing for our papers and thesis work we went through and for her open door anytime I was stuck or needed advice. Her advocacy for women in research is apparent, quite formidable, and inspiring. I'd also like to thank Dr. Amy Pritchett, co-leader of the CEC, as she became more like a co-advisor and consistently challenges my skills as a researcher and software engineer.

Besides my advisor, I'd also like to sincerely thank other members of my committee: Dr. Richard Fujimoto, Dr. Ashok Goel, Dr. Sonia Chernova, and Dr. Terry Fong, for their insightful feedback and encouragement, as well as the difficult questions that confront my perspectives and approaches. Each committee member was someone who had left upon me a remarkable impression from their courses or our discussions with regards to their passion, enthusiasm, and dedication to their students and their work. I'd like to also thank the numerous researchers and friends I've met at NASA AMES: Tamar Cohen, Massimo Vespignani, David Lees, Mark Micire, Brian Colten, Jonathan Bruce, Antoine Tardy, Jessica Marquez, and Yunkyung Kim. I'd like to thank Terry in particular for sparking my initial interest in this work and being a wonderful mentor and guide at NASA Ames and after. While attentive and detail orientated, he is also able to question me with the bigger question in mind, allowing me to be constantly challenged and better for it. I could not imagine this work without our many thought-provoking interactions.

I thank my fellow labmates for all the entertaining discussions, food, and fun these past few years. They were integral to making me feel welcome and making my first few years at Tech especially memorable. I'd like to thank the generation before me: Matthew Miller and

Marc Canelles for being inspirations, friends, and mentors to me. Thanks also to Rachel Haga for being wonderful and there in every and all circumstances; I appreciate how our discussions can switch from memes to research in a snap. I'd also like to thank my peers in the lab, Raunak Battacharyya and Martijn IJtsma, for being patient with me, always challenging my ideas, and working tirelessly alongside me to fix our bugs on our project. My labmates were my home away from home and we've built a camaraderie through all our shared experiences these years through all the projects, classes, quals, exams, deadlines, and much more. I am especially thankful that these friendships will persist beyond our short (haha) time at Tech.

Next, I am grateful for old friends and newly found ones at Tech and in Atlanta that have shaped my experience living and working away from my comfort zone. I'd like to thank Lyes Khalil for being so responsible regarding deadlines and timelines, for being so thoughtful during our research discussions, and for believing in me in every endeavor I sought despite myself not always being the most confident. I'd like to also thank other friends: Antoinette, Patrick, Rakshit, Caleb, Harsh, Karl, and many more for all their goodwill these past years and helping me find ways to celebrate the good and bad. Thanks also to my other friends: Lydia Han for always helping me relax and listening, Amanda Chin for our introspective conversations, Jeannie Wu for always being an open ear, Sophia Han for being the smallest big sister I never had, Alice Chen for being the most positive and adult, Daniel Du for his 4D personality, Stephanie Tung and Ray Kyaw for always accommodating me, Ivan Zhou for the tough conversations, Jenny Jung for introducing me to so much food and her support, Sam Hong for all the rides and talks, as well as countless others: Carson Tang, Anhang Zhu, Jenny Mi, William Yan, Alan Choi, James Ray, Caroline Lee, Grace Choi, Sharon Lin, Chris Xie, Kevin DiPasupil, and Morgan Gaines. I'd also like to thank Edward Wu for his unconditional love and encouragement through the ups and downs of this Ph.D. journey.

Last but not least, I would like to express my gratitude towards my mother and father for

their endless support of me and my education in every way from elementary to my graduate degree. My parents instilled in me from a young age to value education and working hard to achieve your goals no matter how challenging the path there might be. They always have my best interest at heart no matter the circumstance and I cannot express just how incredibly fortunate I am to have them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Human-robot teaming is a complex design trade space with dynamic aspects and particulars. In order to support future day human-robot teams and scenarios, we need to assist team designers and evaluators in understanding core teaming components. This work is centered around teams that complete space missions and operations.

The central scope and theme of this work target the way users should design, evaluate, and think about human-robot teams. This work attempts to do so by defining a framework, conceptual methodology, and operationalized metrics for human-robot teams. We begin by scoping and distilling common components from human-only teaming and human-robot teaming research based in areas such as human factors, cognitive psychology, robotics, and human-robot interaction. Taking these constructs, we derive a framework that describes and organizes the factors, as well as relationships between them. I also present a theoretical methodology to support designers to understand the impact teaming components have on expected interaction. This methodology is implemented for four case studies of distinct team types and scenarios including moving furniture, a SWAT team operation, a rover recon, and an in-orbit maintenance mission. Afterward, we assess various existing methodologies and perspectives to derive metrics operationalized from work allocation.

To test these learnings, I modeled and simulated human-robot teams in action, specifically in an in-orbit maintenance scenario. In addition to analyzing simulation results given different team configurations, task allocations, and teamwork modes, a HITL experiment confirmed a human perspective of robotic team members. This experiment also refines the modeling of teams and validates our performance metrics.

The goal of this work is to provide readers with an understanding that HR teaming can still be greatly expanded. Interaction can be greatly impacted by the definition of components in the teaming framework and also the metrics used to characterize success. This work demonstrates the sensitivity of teams to both teamwork and taskwork.

CHAPTER 1

INTRODUCTION

This research focuses on developing metrics and a framework to better understand human-robot teaming design and evaluation. We will explore best practices through investigating key components of human-robot teaming, modeling and simulating human-robot teams, running an empirical study of these simulated teams, and measuring resulting human-robot team successes and failures. My goal is to not only help designers identify the factors that affect teaming and consideration they need to make for assembling teams, but also create a means to better understand teamwork data to inform them of relationships between interactions and metrics and approximate the impact different teaming interactions have on team performance overall.

This dissertation provides an understand the key components of human-robot team structure, interaction, and teamwork, as well as a method to effectively evaluate these components in testing early in teamwork design to support teaming [1]. Specifically, computational simulation, analysis, and visualizations of results provide a means of understanding the implications of design choices such as team composition, robotic abilities, work allocation, and teamwork mode have on metrics for both teamwork and taskwork performance efficiency within a team. The methods and metrics described here are ideal for robotic designers or mission designers to better understand the trade space they are designing for.

This work will utilize space mission scenarios as these human-robot teams have more inherent constraints and specific taskwork, which are optimal as scenarios for simulation. Longer and more complex missions have a greater risk for error, require high-performing teaming, and more autonomy from Earth as telecommunications become longer [2]. Poor team design can have short and long-term impacts on quantitative mission success metrics, but compound on internal teamwork and team interaction failures. Well designed teams require not just consideration for the input, team members, but also the output, measurable metrics, output and teamwork.

1.1 Research Scope

While we describe human-robot teaming throughout this work, we must scope the conditions of where this work is applicable. The scope of this thesis is on human-robot teams, where robots are fully embodied agents capable of either directed or autonomous actions. The basis of this work is based in human-human teaming research. This work does not apply directly to human-human teams, as it does not span the full set of components known to such teams. We ignore human-human only traits along with other components of humans such as social measures and norms, emotions, feelings, and fatigue. Additionally, this work assumes that robots have some physical embodiment which humans can interact with. This implies that there is some level of physical interaction that needs be addressed along with the more remote and communicative interaction that occurs in interface-only agents. As such, this work has limitations on its applications to purely computational robotic teammates.

The capabilities and descriptions of these types of robots are more or less limited by current-day robotic builds. Specifically, we gear towards trends in robotics within space applications. The use of artificial intelligence or machine learning with these robotics are applicable if the robotic systems comply with the earlier restrictions.

1.2 Requirements of Human-Robot Team Design

As stated, the goal of this work is to assist designers. we need to first address their needs. It is useful to outline the broad set of questions designers need to answer when designing teams with robotic systems.

Team Design Considerations

How do we begin thinking about building a team to performa a given set of tasks?

What aspects of teams have the biggest impact on their performance?

What compositions of team members are best?

What kinds of interaction mechanisms are available to use?

Which interaction mechanisms work best and under what circumstances?

How to measure teams

What are metrics to consider beyond the typical taskwork measurements of performance and efficiency?

How can we compare various team configurations quickly?

Success Criteria and Methods

How do quantitative, objective measures compare to qualitative, subjective measures of taskwork in human perceptions of these teams?

How much impact does teamwork have on the team and its performance?

How can we compare tradeoffs between teams?

1.3 Research Questions and Objectives

Through our exploration of scope and team designer concerns, we find one major objective for research in HRI is the development of methods to evaluate and understand elements that influence HR teaming. With the goal to create and measure human-robot teams as evidenced by mission success and internal health. One big picture question:

How should designers design, evaluate, and think about Human-Robot Teams?

To address this overarching research question, I define more specific questions that answer each aspect of the broader question at large.

- What are the key components of human-robot teaming; how do they relate to one another and influence the design trade space?
- How can human-robot teaming be measured and evaluated fully? What are the significant metrics and measures that are impacted by earlier design decisions?
- How do different design considerations impact the human perspective of human-robot teams? How in turn should these perspectives better team health and performance?

1.4 Technical Approach

The main research goal is focused on understanding how team design decisions may impact human-robot teaming, specifically through the resulting interactions effect on internal and external team components and metrics. To do so, the proposed thesis will implement several phases, see Table 1.1. We will first discuss and survey teaming literature (human-human, human-system, and human-robot) to determine the core components and metrics for those components within human-robot teaming. Next, using this knowledge, we will demonstrate how to translate these concepts into a computational model and simulate human-robot teams (actions, interactions). Specifically, we will present a case study to illustrate the method. With the raw data from the simulation, I analyze how various interaction changes can effect teaming metrics and outcomes. Finally, this work will conclude with an empirical HITL study to validate the modeling and simulation.

This thesis work will answer the above questions through defining human-robot teaming (HRT) amongst various research areas that discuss teamwork, key interaction patterns between human and robot team members, key measurable metrics, and developing an analysis tool for designers to further evaluate human-robot teams. The scope of human-robot teaming is within HRI but derives specific schools of thought from human factors, cognitive engineering combined with the technical advantages of simulation, data visualization, and analysis. The modeling and simulations of human-robot teams uses and advances Work Models that Compute (WMC), a computational simulation framework.

Table 1.1: Research Method

Step	Objective
Foundations and Metrics	
Identifies key human-robot teaming concepts to consider for team designers	Scopes human-robot teaming and defines a framework and methodology for understanding impactful human-robot teaming components and interaction. Identifies metrics for human-robot teaming that span external and internal team health and success.
Modeling and Simulation	
Defines the models and simulations for human-robot teams, with varying metrics, failures, and teamwork actions.	Investigates various robots and capabilities, and failures in robots and teaming. Models measurable metrics and interactions in simulation case study
Exploration and Analysis	
Uses teamwork models to investigate how interactions for HR teaming affect teaming metrics.	Describes graph network and data visualizations to evaluate shifting teaming components and metrics.
Verification and Validation	
The design and results of a HITL study based on previous scenarios of human-robot teams.	Questionnaire results that verify the model of teamwork and interaction described above Participant data analysis comparison to WMC scenario results to validate our modeled metrics.

1.5 Theoretical and Practical Significance

Various types and combinations of human-robot teams will result in different types of interactions between humans and robot counterparts. These interactions will require designers to thoughtfully assemble, monitor, and evaluate their teams in order to understand the ongoing teamwork and team health. Team designers and robot designers alike will need to understand various factors beyond the taskwork and team composition and should be cognizant of not just team performance externally, but teamwork and internal health of the team to create purposeful human-robot teams.

This work will create a comprehensive study of the current and future state of human-robot teaming by 1) re-iterating key components of human-robot teaming and focusing on interactions and metrics 2) verifying and validating correct modeling of these teaming components, metrics, and simulating multiple scenarios and missions, and 3) developing a data visualization toolkit to help team designers evaluate design decisions on human-robot teams early-in design.

These tools will help team designers investigate various research areas understanding of teamwork, allow them to forecast predictions, and test their hypothesis on different measures of effectiveness and early indicators of success in their HR teams. My goal is to introduce computational analysis methods of computational simulation and data visualization to evaluate the breadth and depth of HR teaming. This will provide a means for designers to understand their teams and utilize them to their full potential, be the best fit for the mission, preparedness, missions success.

1.5.1 Contributions

This thesis will provide the following contributions to HRI research.

1. Created a new comprehensive framework for human-robot teaming by combining the key components of team design and interaction.

2. Developed a method to identify distinct archetypes of interaction in human-robot teams (and showed how they fit into a universal framework).

3. Derived metrics from the HRT framework to capture the teaming elements beyond performance and efficiency. Operationalized the method and metrics in a computational framework for simulation and analysis.

4. Extended existing computational framework for function allocation to include the metrics.

5. Demonstrated the sensitivity of effective teams to attributes of both teamwork and taskwork.

The following chapters will describe my approach to investigating teaming from the bottom up: beginning from understanding what a team and teaming is, to key teaming components and metrics, to modeling these components and simulating different teams working together, to the visual analysis and breakdown of these teams, to validation of the technical modeling and simulation with a HITL.

CHAPTER 2

HUMAN-ROBOT TEAMING CONCEPTUAL FRAMEWORK

Future human-robot teams will require effective teamwork for mission success as robotic capabilities advance and mission objectives become more complex. Designing for human-robot teams goes beyond accounting for specific factors and beyond to evaluating entire processes of teaming and teamwork. Understanding these components requires a holistic consideration of team performance, interaction, and success. Team designers and system evaluators should, therefore, be knowledgeable of both the defining inputs and key outputs of teaming that will affect design decisions as well as the overall effectiveness of the team. However, teaming and interaction shift dynamically throughout a mission or a series of tasks; there are many challenges in determining what is a measure of team success as well as the multitude and layers of complexity that factor into team success. To address these challenges, there is a clear need to define human-robot teaming and its components more clearly.

While human-robot teams have been explored by many research areas, a comprehensive evaluation of key factors to consider for team interaction remains largely undefined. Human-robot teaming, in particular, is relatively new compared to the study of human-only teams. From human factors to cognitive psychology to robotics and HRI and more, the concept of teaming has been refined to the specific viewpoint of each respective area. Collectively, these distinct disciplines have investigated various types of teams, specifically human-robot teaming, and made valuable contributions that provide context to understanding them at distinctive levels of abstraction.

Human-robot teaming shares a subset of various characteristics from preceding research regarding other types of teaming: human-human, human-animal, human-automation teaming, and thus can begin to compile a theoretical foundation from these works. This paper describes the basis of human-robot teams through takeaways from prior research that highlight a multitude of complex components and their relationships. Through analyzing these different areas, we gather the main components of teaming to be *internal*: team members/individuals, team structure, and team interaction, and *external*: taskwork and external factors(to the team). We continue to break down each of these components to adapt them into a general conceptual framework of teamwork and interaction that illustrates associations and dependencies between these components. This framework of teaming follows an integrated description of teamwork: a group of individuals [humans and robots] working together, effectively, towards a shared goal. Additionally, we explore Interaction as a central theme to human-robot teaming and dive further into centering the framework around interaction as the main component of teaming. To show examples of applying this conceptual framework to distinct teams, we present four distinct human-robot teams and evaluate the meta-interaction of these teaming components with each other. Specifically, we show how our teamwork framework can apply across teams with various changes in team structure, team members, taskwork, and external factors.

This work is motivated to help other team designers and evaluators how to better understand human-robot teaming and teamwork. We hope to present a series of components that inspire team designers to not only investigate teaming more comprehensively but also remind researchers of forgotten components that are important to define, measure, or be aware of. This work also serves to provide a basis for components of teaming that future work can continue to build upon. Our framework of human-robot teaming will serve to illustrate the various concepts that designers should keep in mind to help support creating and understanding teams.

2.1 A Review of Teaming Definitions Across Disciplines

Prior research for teaming in general spans across multiple areas, including Cognitive Psychology, Human Factors, Human-Robot Interaction, Robotics, and more. Building the bedrock for human-robot teaming starts with rooting this topic among the other areas and defining with clarity what human-robot teaming encompasses as an area, as well as the important considerations to be made for scoping. Human-robot teaming can be broad as it applies to any number of team members or systems working in an environment. Here, we will refer to it as a combination of human(s) and robot(s) working together on a task, effectively towards a shared goal [3].

Human-robot teaming has been previously evaluated with the perspective of robots acting as 'tools' while humans members typically manually operate or control them [4]. This view when applied out of context keeps robots as interfaces in a mission without consideration as a 'team member'. This limitation can impact performance due to dependencies on human training, skills, and more. For example, robots that require constant monitoring, attention, or human skill to function (a robotic arm) can cause high workload stress on human counterparts and can lead to degrading situation awareness. Inflexible systems like these ultimately reduce the type of work that these human-robot teams can perform though they may be useful for specific use cases.

Another, more recent, perspective of human-robot teaming considers robots more-so as 'team members' [4, 5, 6]. In this case, humans and robots are interdependent and can share various roles or responsibilities throughout missions. These relationships can span the spectrum from supervisor-subordinate to peer-peer. While this point of view is still gaining momentum, declaring robots as 'team members' requires considerably more scrutiny to engender 'teamwork' as we understand it between human relationships and teams.

2.1.1 Human-Human Teaming

Human teaming has been studied extensively in the Cognitive/Organizational Psychology and Human Factors literature. As a result, a number of relevant concepts have been identified and studied, and frameworks have been developed to encompass the various dimensions of teaming. In this section, we do not attempt to review the entire literature (see [7, 8, 9] for comprehensive surveys), but rather we focus on the core perspectives of human teaming and some of the key concepts that have come out of the Psychological literature.

A *team* can be defined as "two or more individuals with specified roles interacting adaptively, interdependently, and dynamically toward a common and valued goal" [9]. Teams can be either *co-located*, in which all the agents are physically co-present, or *distributed*, in which the agents are spread across multiple spatial ranges and/or timescales. We will focus our review in this area on work teams (rather than, say, sports teams or social organizations). Work teams exist to perform an interdependent task typically involving social interaction, communication, and shared goals [8]. According to [10], there are six types of work teams that have been identified, including management, service, production, project, action, and parallel teams. However, since it has been argued that team processes and the task itself are more relevant for understanding team effectiveness, we will focus on those elements in our review rather than attempting to distinguish between different types of teams. Regarding team effectiveness, it is often conflated with team performance. *Team performance* can be viewed as a subset of team effectiveness, involving only performance on a task, or the outcome. However, *team effectiveness* captures additional aspects of the interaction that may or may not affect performance. For example, a team can perform very well on a task but display poor communication, low confidence, and limited adaptability.

Over the years, various frameworks have been proposed to capture the relevant processes involved in effective human teaming. The McGrath framework for team effectiveness was among the earliest approaches, and it has been influential in shaping our understanding of the processes involved in team performance and effectiveness [11]. The framework emphasizes the central components of teaming, which include inputs, processes, and outputs; as a result, it has been called the *IPO* model. The IPO framework involves *Inputs*, which are constraints on team interaction imposed by various aspects of the team/task, including the organizational structure of the team, the nature of the task, and environmental constraints. These inputs feed into team *Processes*, which describe how these inputs influence team interaction with regards to carrying out the task. *Output* is the result of the team activity, and can include measures of performance as well as attitudes (internal and external) about the team. This framework has been extended over the years to add environmental constraints, interaction over time, and nested structures within each level. As it became evident that not all factors that influence team outcomes can be considered "processes", the framework was extended to *Input-Mediator-Output-Input (IMOI)* by [12] (see Fig. 2.1). *Mediators* include processes such as coordination, feedback, and planning, which are distinguished from *emergent states* such as team confidence, climate, and cohesion. Mediators can exist at the individual level, the group level, and the organizational level, with different effects at each of these levels. The additional "Input" at the end of IMOI represents the cyclical nature of the framework and the dynamic feedback between levels.

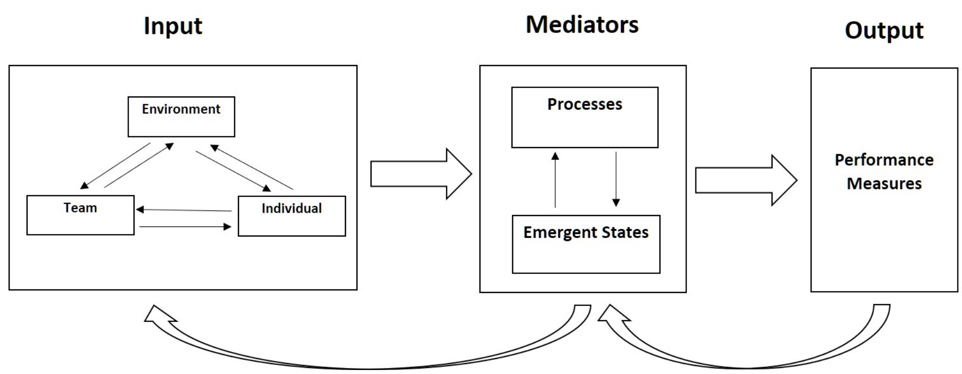


Figure 2.1: Input-Mediator-Output-Input (IMOI) Framework. Boxes represent distinct concepts/factors and arrows indicate causal influence.

Team Mediators

It is important to discuss these mediators in some depth, as they represent the critical components of teaming that must be part of any comprehensive account of human teaming. These mediators include team processes and emergent states.

Team Processes are split in terms of *taskwork* (strategies, operating procedures, etc.) and *teamwork* (interaction between team members). Three categories of processes have been identified by [13] in their taxonomy: *transition*, *action*, and *interpersonal*. Transition processes represent the early stages of teaming in which the goals are formulated and strategies/plans are developed to achieve these goals. These include processes such as mission planning, goal specification, and strategy formulation. Action processes represent the performance of the actual task activities as well as coordinating with teammates and monitoring their status; these include monitoring task progress, tracking resources, assisting teammates, and coordination. Finally, interpersonal processes involve aspects related to trust, motivation, and interpersonal dynamics (e.g., conflict management, motivation building, etc.). There is evidence that these factors can both positively [14] and negatively [15] impact team performance.

In addition to team processes, a number of *emergent states* have been identified which have bearing on teamwork. These emergent states are different from processes in that they are dynamic properties of the team that are related to cognitive and affective states of the individuals rather than interaction processes of the team as a whole. [8] have compiled a list of the most studied emergent states, which include the following. *Team Confidence* involves beliefs about competence as it relates to a specific task (efficacy) or more general confidence (potency) shared by all team members [7]. These states have been correlated with improved team performance [16]. *Team Empowerment* distinguishes between structural and psychological empowerment. Structural empowerment deals with how the structure of the team can lead to effects on performance and satisfaction, whereas psychological empowerment deals with either an individual or collective sense of authority in terms of controlling outcomes. Psychological empowerment is associated with improvements in performance and customer satisfaction [17]. *Climate* is a team-level state which has to do with the attitudes and expectations by which the team operates. Climate has multiple dimensions and has been broken up into safety, service, and justice climate. *Safety Climate* refers to team attitudes about safety procedures, and perceptions of safety climate have been associated with the rate of accidents and injury [18]. *Service Climate* refers to organization-level concerns about customer satisfaction, and is associated with perceptions of service quality [19]. Finally, *Justice Climate* refers to a team-level sentiment about how the team is treated, and it is associated with performance, commitment, and absenteeism [20]. *Cohesion* is a team-level state which emerges somewhat late in a group's operation, and broadly deals with the unity or bond that is formed in the team. Cohesion has been decomposed into three component categories of interpersonal attraction, task commitment, and group pride, and has been strongly associated with effective team performance [21]. *Trust* broadly describes the belief that team members will fulfill their part of the task without the need to monitor or intervene [22]. *Shared Mental Models (SMMs)* are structured representations of various team- and task-related components that are shared among all teammates, and that are essential for effective teaming. SMMs include information about the task itself (*Task-SMM*) such as the environment, the required equipment, the operating procedures, strategies, etc. They also include information about the team (*Team-SMM*) such as the various roles, interdependencies, communication channels, and mental models of the individual teammates (their beliefs, goals, intentions, etc.). There is evidence that the team SMM directly impacts performance, whereas the task SMM has a more indirect effect [23]. Finally, *Strategic consensus* describes the strategic priorities shared among team leaders. Whereas an SMM is shared among all team members, strategic consensus may only be available to managers or other individuals towards the top of the team hierarchy [24].

Measuring these team mediators as a function of team effectiveness has been a particular challenge over the years. One approach comes from [25] who developed three main categories to measure team outcomes/effectiveness, including *performance*, *attitudes*, and *behavior*. Performance can be further broken down into organizational-level performance, team performance behaviors and outcomes, and role-based performance [8]. Each of these categories comes with a host of possible outcome measures, including composite measures ([26, 27]. Regarding attitudes about the team, the prominent research directions have been in *team, job, and organizational satisfaction* and *team and organizational commitment* [8].

Team viability is another related construct and refers to a team's potential to stay together [28]. In general, team effectiveness should be measured in a way that corresponds with the particular team. In a team whose goal is to produce a specific outcome, the quality of that outcome should be the performance measure. In a team that involves customer service, customer satisfaction should be the main outcome measure.

Core Components of Teaming: "The Big Five"

Due to the sheer number of mediators that have been implicated in effective teaming, it is difficult to practically apply findings from the literature. To address this issue, there have been some attempts to focus on only a core subset of the mediators in order to capture the essential requirements of teaming. One such attempt comes from [9], who proposed five key attributes that are vital for effective teamwork, and which can serve as a framework to describe the key dimensions or requirements of teaming. These attributes include team leadership, mutual performance monitoring, backup behavior, adaptability, and team orientation. *Team Leadership* involves one or more agents on a team that are responsible for several executive functions. First, they define goals, manage resources, assign roles, and perform other leadership tasks to ensure the functioning of the team. Transition processes [13] which involve the early stages of team formulation and planning are typically carried out by the leader(s). In general, these responsibilities can be viewed as establishing and maintaining the team's SMM. Leaders are also responsible for adapting the team in the face of changes and managing conflict by establishing behavioral norms which support all the other components of team effectiveness. *Mutual performance monitoring* involves the mutual monitoring and tracking of the tasks and performance of other agents on the team. Mutual performance monitoring requires an understanding of the task as well as the role of each agent in accomplishing the task, i.e., an SMM. An open, friendly, team climate is also a critical requirement, which ensures that agents view this mutual monitoring in a positive light. *Backup behavior* involves assisting another teammate who cannot perform their task effectively due to, e.g., being under high workload. The assistance can come in the form of giving advice/coaching, assisting in the task, or completing the task for them. These behaviors are important in highly interdependent tasks since the delay or failure in one subtask can negatively impact the shared team task. Mutual performance monitoring is a requirement and allows for assessment of a teammate's workload. *Adaptability* involves recognizing changes or issues as they arise and re-adjusting the team to compensate. The issues can be internal to the team (e.g., conflict, failure, etc.) or external (e.g., environmental change, contingency, etc.). This requires all of the components described above, including maintaining an SMM to anticipate how the new changes will affect the team/task, monitoring teammates to determine if adaptation is needed, backup behavior to assist and accommodate as necessary, and leadership to initiate and manage these adaptations. Finally, *team orientation* refers to teammates maintaining a favorable, team-oriented attitude during the task. It has been shown to improve various aspects of team effectiveness, including decision-making, coordination, and satisfaction. Team orientation has its effect due to the ways in which it supports the other components. For example, teammates are more likely to monitor and assist one another when there is a general attitude of cooperation.

In addition to these Big 5 attributes, [9] propose three coordination mechanisms which are additional requirements needed to support the five critical components of teaming. These include SMMs, closed-loop communication, and mutual trust. SMMs, as defined above, describe organized representations of team and task knowledge that are shared and updated among teammates [23]. SMMs are considered a primary coordination mechanism because they enforce a shared understanding among teammates and ensure that the team has updated information about the critical components that are needed to facilitate joint behavior. In this way, SMMs directly enable the Big 5 attributes above. *Closed Loop Communication* emphasizes that communication in teams is critical for coordinating joint activity. Of course there are tasks in which communication is limited or even absent, but generally, work teams performing interdependent tasks require some degree of communication between teammates. However, communication alone is not enough since real-world task environments are often noisy, and the message may not get through. Thus, it is important that the communication is “closed-loop”, meaning that not only should the listener acknowledge the message, but the speaker should also confirm that the message was correctly interpreted. This is similar to models of common ground from the Psycholinguistics literature in which a discourse “contribution” is an exchange in which mutual belief has been reached by all parties about the meaning of the exchange [29]. *Mutual Trust* is defined as an understanding that teammates will perform their part of the task in a timely and appropriate manner. It has been shown to support various aspects of teaming, including confidence in teammates, help-seeking behavior, and free exchange of information - all of which are associated with improved cooperation and teamwork outcomes [30].

Overall, these Big 5 attributes highlight some of the key requirements for effective human teams, and the coordination mechanisms describe additional requirements which are necessary to support the critical components of teaming. Due to the inherent complexity in the teaming literature, frameworks such as this are extremely useful to help researchers focus on the key attributes that facilitate team effectiveness.

2.1.2 Human-Agent Teaming

As robotics technology advanced, Human-Machine Interaction (HMI) formed as new team formations included humans and automation together. While these human-machine teams early on often used these robotic systems and interfaces as assistants or tools in mission operations, the concept of these systems acting as team members has now become more of a reality. As such, research that explores the interaction between humans and robotics/automation has grown alongside human-human teaming. We explore two umbrella perspectives of human-agent teaming through Human Factors Cognitive Engineering and HRI’s here.

Cognitive Engineering Perspective from Human Factors

Human Factors is an interdisciplinary field comprised of engineers, psychologists, economists and many others who focus on improving human-automation interaction (HAI) through better design [31]. As a multifaceted area, the foundations of this field stem from cognitive psychology, ergonomics, industrial design, and more, which all come together to provide insight on how these teams might be improved when automation and machines become integrated with humans. Human factors from a big picture perspective help researchers to better comprehend the human's role within human-automation and human-machine teams [31].

Core components of interest identified by human factors have also branched into their own sub-topics. One key area of focus in human factors is understanding the work to be done within a team. Work analysis methods include work domain analysis and traditional task analysis (decision ladders, workflows) [32]. Cognitive Work Analysis is one specific method for break down and understand the workflow and structure of mission taskwork [32]. Work allocation is another area of investigation that is focused on taskload and taskwork assignment to team members, and decision making [33, 34, 35]. Work allocation is the design decision for how to allocation tasks to agents [36, 33]. This concept has been a useful tool in understanding human-system teams and interaction, particularly early in design. The formal decomposition of taskwork and its distribution across agents will set the requirements for teamwork and thus, team interaction [32].

With the knowledge of human impact and influence in HMI/HAI comes the ability to help design teams of humans and automation. Given the insight into the human perspective, Human Factors Cognitive Engineering has evolved into a design heavy research area. When designing joint human-machine and human-automation systems, common questions the Human Factors community aims to answer often involve improving workload and taskwork factors on the humans in these teams. More specifically, researchers may investigate how different distributions of work can impact the overall performance and experience of agents. Specifically, within cognitive engineering, significant efforts have been made to understand workload distribution and function allocation which attempts to identify key taskwork constraints and assignments.

To help address design considerations, human factors has contributed frameworks, metrics, and methologies. Frameworks for cognitive engineering to better think about these HAI problems, identify metrics and concepts of interest to the community, and develop methodologies and practices for analyzing and understanding these human-automation teams. These frameworks help users analyze joint human-machine systems. Key frameworks from this field have helped researchers approach and diagnose these human-machine systems from a school of thought that is human-centric.

Measures of human-automation teams help researchers better understand how to improve human performance and efficiency. This human-centric perspective has led to many key metrics of interest including mental workload mental models, trust in automation, and taskwork distribution [31]. A core driver of these measures is to understand teams and team cognition [32]. To measure human experience, human factors employs metrics that address the causes of human performance and the limitations of it. Previous work in cognitive engineering has focused primarily on performance metrics such as total time to complete tasks, human involvement time, wait or idle time, resource consumption, or task-specific performance measurements [37, 38]. Estimates of workload, based on the number and type of task that the human is performing, are traded off against performance estimates when finding an optimal work allocation [39]. Other work addresses the reliability of work allocation by quantifying and minimizing the probability for failure and human error [40]. Finally, there is a large body of research in aviation and related domains that focus on human performance metrics such as situation awareness and skill degradation [41].

Methodologies from this area focus on both the design and analysis of human-automation interaction and man-machine systems or teams. To investigate team design decisions on human and team performance, human factors research has observed the impact of role allocation and task or function allocation. Post analysis of these teams includes using methodologies such as Cognitive Work Analysis, Cognitive Task Analysis, Goal-Directed Task Analysis, and descriptive understanding [42, 43, 44, 45, 46]. Researchers often start by defining taskwork either through Hierarchical Task Analysis or Work Domain Analysis [44, 47], which decompose work into different conceptual levels. Other methodologies that represent models for HAI processes [31]. These methodologies have been applied to a wide variety of domains fields (aerospace, aircrafts, ATC) to asses taskwork, information flow, and domain constraints. For human-robot teaming, these cognitive models and methods are useful in a deeper analysis of the human's role when working with automation of any kind. Models like Rasmussen's influence in early human factors work is widespread. In particular his Abstraction Hierarchy has been used in various domains to better create human-machine interfaces by typically describing the operator's work domain that also reflects the intentions and goals. The original AH contains the following: functional purpose, abstract function, generalized functions, physical functions, and physical form [48]. Rasmussen's hierarchy of interaction anticipates frameworks and methods for testing how HAI metrics change when varying team inputs [49].

When applied to understanding teams, Human Factors provides key insights into the right role for humans to take on in such teams to pair with advancing automation with respect to teamwork and taskwork [32]. Human factors views HRI as either an extension of HAI or a specialized version of HAI. Specifically, humans factors applied to studying human-robot teams have explored various missions and metrics of interest to team designers. Some of this work has effectively modeled mission work, failures and recovery in robotic systems, teamwork modes between humans and robots, communication, and physical interactions [50, 51, 52]. Metrics under investigation went beyond typical measures of overall mission performance (i.e. mission duration, idle time, taskload) and expanded to understanding relationships between metrics such as failure timing to robotic capabilities and action interdependencies to failure awareness. These results showed the importance of modeling various types of taskwork and interaction in human-robot teams and the usefulness of human factors as a perspective to study them. For human-robot teaming, it is even more important to provide methods for designing technology to maximize joint-human-automation performance while not sacrificing the human's contribution.

Robotics and HRI Perspective

Robotics has commonly investigated robots in teams from a quantitative perspective around a technically bound object. Some of this work extends to telerobotics and teleoperation [31]. In particular, this field including supervisory control has defined multiple levels of automation and its impact in human-machine interaction [53]. As robotics advanced beyond manual control and complexities in interaction grew, Human-Robot Interaction (HRI) became more and more relevant to team success. HRI has investigated teaming and is now trending towards robots as equal teammates instead of merely a physical extension for humans or objects to do unwanted or menial work [54, 1, 2, 55]. Specifically, the HRI problem is: *'to understand and shape the interactions between one or more humans and one or more robots'* [31]. One primary goal of research in this HRI has been to "investigate natural means by which a human can interact and communicate with a robot" [56].

Within HRI, one broad definition of teamwork has been explained by the following life cycle by Kozlowski and Bell including 1) Team Composition, 2) Formation, Socialization, and Development, 3) Effectiveness, Processes, and Enhancements, 4) Team Leadership, Motivation, and 5) Continuance and Decline [3]. Throughout this life cycle, Kozlowski and Bell highlight and link together important concepts and components to teaming that are different levels of abstraction. While all components can be further investigated, the majority of this paper will explore and focus on stage three, four and five of the lifecycle components: effectiveness, processes, enhancements, team leadership and motivation, and continuance and decline.

HRI has also produced many frameworks, models, and methods of interaction to further evaluations of human-robot teaming. Models and frameworks describe HRI from a conceptual perspective, while methodologies are often more specific evaluations of teamwork components.

One model of important in Scholtz's cycle of cognitive interaction internal to each individual [57]. Scholtz's model of interaction is based on an HCI model of internal, cognitive processes. In particular, she defines her framework of interaction between two agents (a human and a robot) through a cycle between goals, intent, actions, perceptions, and evaluations affected by agent roles and situational awareness [57]. Klein highlights a joint action ladder for conversational interaction through attending (noticing), perceiving, understanding, and acting (acknowledge) [58]. This joint action ladder contains similar aspects of interaction as Scholtz, whereby the Scholtz model includes some precursors to the ladder with goals and intent. Cohen and Levesque focus on a formal approach to understanding teamwork in a way that would allow for artificial agents to serve as teammates with humans [59]. They describe agents acting based on their mental states: beliefs, desires, goals, intentions, etc. These agents need to balance their own mental states with that of their teammates in order to manage joint intentions and individual ones. Here, joint intention towards a shared goal is the primary motivator of team activity. Their framework includes individual commitment (persistent goal), individual intention (to achieve a goal), joint commitment to a goal, in which the team works to achieve a common goal and maintain that goal until it is mutually known that the goal has been achieved or some teammate communicates otherwise. Communication is key to establishing common ground after teams diverge on information or beliefs. One open question from this work is when communication is necessary for teamwork. While these models and frameworks focus on different aspects of HRI, they share many similar aspects including the formulation of goals and common ground, the desires, and intentions of the agents, the action they take to complete the taskwork, and their perceptions and understanding of their actions and surroundings.

Many HRI methodologies target specific metrics of interest through HITL experiments and field studies, yet there remains to be a cohesive foundation of methods and practices [56]. While many concepts of HRI have been identified and thoroughly evaluated through such studies, these methodologies are often not broad enough to apply to other teams. One example of a methodology that has been applied to diverse human-robot teams is Johnson's Co-active Design. Co-active Design defines critical teamwork as managing interdependence relationships between agents, where interdependence is highly dependent on the capabilities of agents [60]. Johnson's method, in particular, provides support for interdependent teams to work effectively.

When applied towards teams, HRI addresses big-picture questions like how many robots can a single human manage, how do different combinations or team structures impact mission success, and what aspects of dynamic teaming are important to be aware of and control? Generally, the answers to these questions are complex, requiring understanding the available communication modes, control modes, cognitive load on humans, attention required from humans, task type and feedback, and autonomy of robot among many others [31]. While previous trends were geared towards defining specific robotic capabilities, more researchers are now also testing limits of other combinations of factors outside of the agents in the team. Due to the infinite combinations of these components to be discussed in Section 3, the overall organization of these concepts and their relations to one another is incomplete.

2.1.3 Combining the Perspectives of Teaming

We can see through defining these distinct areas that while each has clearly influenced and discovered important points critical to HRT, much of the vernacular within these respective areas often overlap. Similar concepts that are of major interest appear across multiple disciplines and in many cases are nearly synonymous. Indeed imagining the overlap of these areas brings to mind a complex Venn diagram, each point of view crossing over another with respect to their general point of view by which they approach the study of teams. These findings across different types of teams (human-human, human-automation, multi-agent) can apply to human-robot teaming with some pruning.

One example of shared and synonymous terminology is team "flexibility" vs "adaptability". Both insinuate the ability of the team or individual to react in off-nominal and unpredictable situations and are used in different research areas [61, 62, 63, 64, 65]. Other examples of similar concepts include "recognition" vs "comprehension" vs "reasoning" [66, 67, 68, 69]. Some may argue that comprehension cannot occur without recognition or reasoning, and others may defend that recognition and reasoning is effectively the same as comprehension in practice. Here we see another interesting phenomenon, where in some cases, certain terms are used to define the others. Another example is "trust" and "reliance" [70, 71, 22, 72, 73, 74, 75]. Although trust is a large and complex social behavioral term. Reliance could also fit under trust as an umbrella term. "Common ground", "shared cognition" and "shared mental model" are also seemingly related, as the first two are also considered the result of forming the latter [29, 76, 77, 23, 78, 79, 80, 81]. However, while we do find that often times multiple researchers are referring to the same core concept, the specific definition or connotation of synonymous terms tends to slightly vary.

Interestingly, while there are many terms that are synonymous and therefore have the same concept, there are also terms that are shared across research areas that differ in their definitions. In this case, a particular aspect of the term and even what perspective is being studied by differentiates these research areas and findings. Some terms like communication, which is broadly referred to in cognitive and behavioral psychology, can be referred to by others as information transfer or even interfacing. Communication has also been discussed in terms of the intent of the user, and whether or not it should only be considered if the recipient receives the information. Typically, the muddling of definitions for distinct terms occurs in more abstract and vague concepts. Another example is a lack of distinction between "collaboration", "cooperation", and "coordination" [63, 66, 64, 82, 5, 70]. These terms have loosely defined in the literature before, although now there are more attempts to clearly define them [50]. Now, we see if we keep pulling this thread, the entanglement between the relationships and dependencies these components share (or do not share) quickly grows in complexity. However, it is clear that a harmonious categorization of the patterns that compose human-robot teaming can be useful to any individual working with or on these teams.

While undoubtedly complicated, human-robot teaming can be situated within all these aforementioned research areas by taking in the components that are integral to teaming while removing concepts that do not fit the underlying dynamic between humans and robots. Thus, while human-human teaming has been thoroughly researched, one question for human-robot team designers is how to leverage this knowledge to teams with robotic members [83]. One of the bigger challenges to shifting human-human teaming to human-robot is that humans and robots have a wide but distinct array of abilities and limitations, be they physical or cognitive [31]. Accounting for the clear differences between the two agents and their relationships with one another is integral to building great teams. To truly understand these new types of relationships, researchers must be aware of the multitude of concepts surrounding the study of teams and any resulting interdependencies [60].

It is important to note that while combining these components from prior teaming research, we must also be aware of components that do not apply to human-robot teaming, particularly from the behavioral literature. For example, many terms that focus on cognitive processes and states are unique to humans. While this is not to say that robots or automation may one day achieve some degree of these cognitive processes, we must bound the overview of terms within reality in the near future to remain relevant. Consider another example with team orientation and mutual trust as attributes that a robot can be imbued with (e.g., a rule to always trust its teammates or to always progress in the task). Therefore, conflict management strategies that deal with interpersonal conflict may not be necessary since a robot will never violate such rules in the first place (so long as its a part of its programming). On the other hand, robots might make different kinds of errors such as misunderstanding a command, inappropriately changing tasks, or just failing entirely. In these cases, it is important that the robot has mechanisms to recover from these failures, and that the humans are aware of these possibilities, as well as strategies for dealing with them.

Based on the findings from multiple research perspectives we can focus on the main aspects: being effective, working together, and having a shared goal. We suggest and operate throughout rest of this paper with, a description of human-robot teaming: ”a group of humans and robots working together, effectively towards a shared goal”. We specify the result of teamwork should embed these three qualifications.

2.2 A Conceptual Framework of Human-Robot Teaming

We present here a conceptual framework of human-robot teaming that brings together the perspectives previously described in section 2.1 and organize the core factors that affect said teams. To dive deeper into the underpinnings of HRT, we need to identify the core components of teaming, i.e. the factors that are effected and affect it. We know already that human-robot teaming is complex and we first segregate the factors that impact it into two areas: internal vs external components to the team. Internal concepts refer to any aspect of teams that concerns team members and their capabilities, team structure, and team interaction. External components of teaming include taskwork and factors external to the team or its members. Each component is explained in detail below:

- **Taskwork**

  The total workload or mission to be completed by the team.

- **External Factors**

  Elements that may be unplanned or unknown and out of the control of team designers. These may be predictable to some degree.

- **Team Members**

The individual traits, abilities, and features of members in a team.

- **Team Structure**

  The team composition, roles, and hierarchy. The organization of the team members and operations.

- **Team Interaction**

  The direct or indirect involvement, influence, and action between team members [31].

We propose these five concepts as the foundation of an HRT framework: the taskwork to complete, the external factors that may affect teaming, the team members' individual abilities, the team's structure, and the interaction between different teammates. Many of the distinct terminology used to describe teams can be filtered into these canopy categories. At a higher level, it should be noted that these concepts have their own specific relationships between them. These relations are suggested in our framework to follow our description of teaming. This framework is presented in Figure 2.2 and shows how the shared goal is central to all other components, able to be updated and changed by each factor. The outer circle which encompasses all the factors represents the ideal outcome of teaming as working together, effectively, towards that shared goal.

While this figure demonstrates the broad, high-level impact and factors for human-robot teaming, it is the lower level relationships and intricacies of these components that will provide an extensive understanding of these terms and their inter-relationships. In the following sections, we will describe first the fully expanded view of human-robot teaming by breaking down each of these five components into the main sub-concepts they are comprised of. After we will dive into the sub-components and branches of these five concepts of our HRT framework.

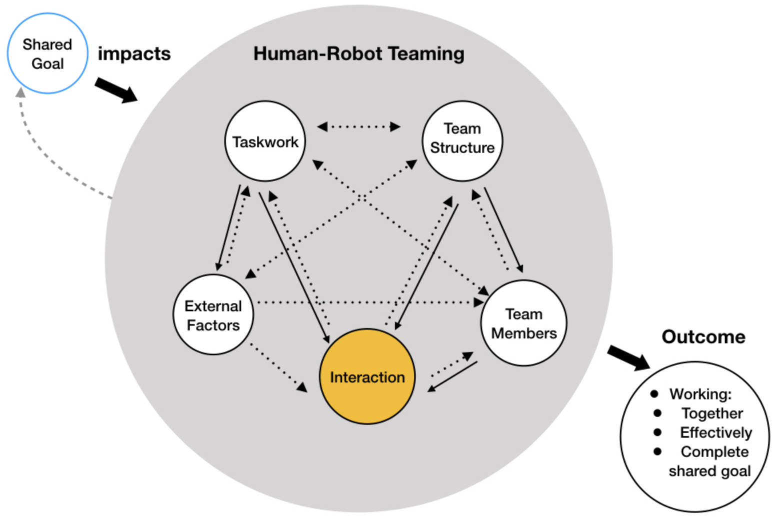


Figure 2.2: Diagram of HRT Framework Structure and Relationships. Dependencies between concepts are represented through solid arrows that imply direct manipulation and dashed arrows that represent the potential to impact other concepts

2.2.1 An Expanded View of Human-Robot Teaming

Before we dive into each component, we can first explore a holistic view of these sub-components and their relationships in Figure 2.3. Here, we can see a variety of detailed concepts that researchers have studied. Interaction is slightly different than the others as it is not divided into sub-components, but rather through a model of the cognitive processes behind interaction [84]. It is important to note this distinction in order to further categorize the interaction terminology under this model for organization.

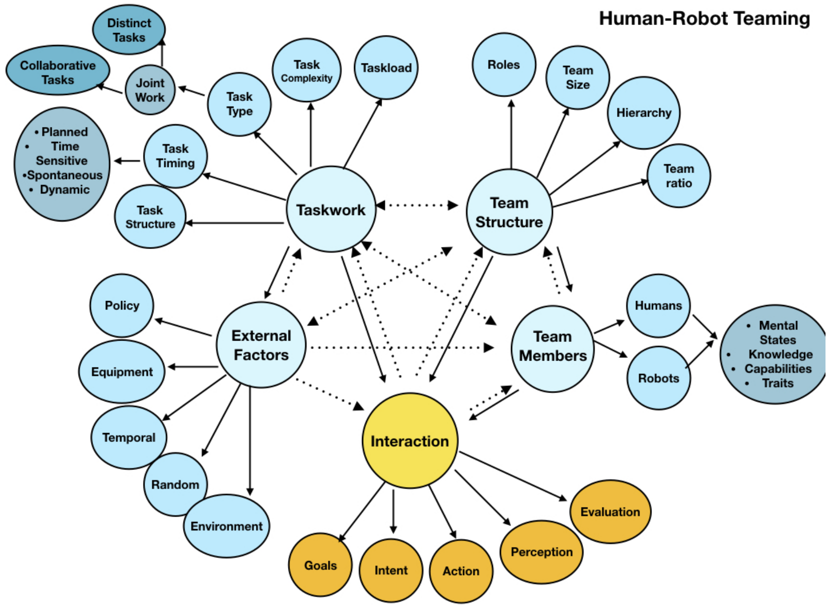


Figure 2.3: Expanded view of each component of teaming, taskwork, external factors, team structure, team members, and interaction. Interaction is slightly different than the others as it is not divided into sub-components, but rather through a model of the cognitive processes behind interaction.

The relationship between how these five components affect each other is implied through the arrows. While arrows between expanded components are not shown, it is implied that certain sub-components can impact each other. These dependencies between concepts are represented through solid arrows that imply direct manipulation and dashed arrows that represent the potential to impact other concepts. That is to say that solid arrows imply changes and the definition of the source will definitively impact the target. For example, lets us examine the solid line from team members to interaction. The individuals that compose a team likely share some commonalities and abilities. However, team individuals can also have distinct traits and capabilities that not only differ between humans and robots but also between human to human (Astronaut vs MCC specialist) and robot to robot (RMS vs. Humanoid). Therefore, team interactions, which is dependent on the two or more individuals interacting, in turn, will be affected by **who** or **what** in the team is interacting. Thus not only the abilities among the team members but also taskwork requirements and team structure directly impact interaction as well (solid line). We review the details of these relationships from the perspective of the source affecting the target concept in question below:

Taskwork nature and placement can definitively change external factors (environment, weather, etc), as the location of taskwork will minimize of maximize the impact these random factors may have on the team. Mission location and timing determines what kinds of external factors to expect). Similarly, taskwork can change the interaction since task timing, structure, taskload, and more can impact what kinds of interaction will arise. Taskwork has the potential ability to impact team structure or team members (i.e. may require agents with specific skill sets).

External Factors may impact all other components, but this is dependent on the type of external factor that the team may or may not deal with. While external factors may certainly change any interaction that can occur (randomness, sudden changes), there may also be cases where there are little to no external factors (highly controlled environments). With perfect conditions or luck, external factors may not play a factor. However, with more uncertain external factors, taskwork, team structure, team members and interaction may all be impacted. Thus we represent the potential that external factors may not have an absolute impact on interaction and other components with a dashed line.

Team Members definitively impacts Interaction, as the combination of different team members will change the relationships of the team and any resulting interaction that would occur as the team executes. The definition of team members does not always impact team structure (all robotic agents of the same type). However in some cases the traits and capabilities of individuals change the chosen structural organization (flexible agents who can take on any role).

Team Structure definitively impacts the Interaction, as the agent roles, team size and hierarchy can change how agents behave towards one another. Team Structure will also impact team members (organizational hierarchy, a chain of command). Team Structure may or may not impact taskwork, or external factors depending on the structure, ratio, and size.

Interaction can be considered as a potential impact factor to the other teaming components. In fact, it is rather more highly impacted by each of the other factors. For example, a change in weather may change how quickly agents must work to finish. A flexible team structure may offer different interdependencies than a rigid team structure. However, depending on circumstances it has the potential to impact any of the other factors. For instance, the adaptability of the team may allow for changes in the types of taskwork required to complete the task. In this case, designers may care more about accomplishing the task regardless of how it is done.

Through this discussion, we can already see the uncertain and situation dependent considerations designers may want to keep in mind. Given this big picture of the conceptual framework of HRT, we can now describe the components of each area in greater detail. Teaming is highly reflective of the context surrounding these five factors. One team may require a different structure, interaction, and taskwork than another. Going forward we stress that teaming is not defined by solely these internal or external components, but instead by how these different concepts interact and affect each other.

2.2.2 Team and Mission Taskwork

In the behavioral literature on human-human teaming, taskwork is defined as a " team's interactions with tasks, tools, machines, and systems" [85]. Generally, this refers to the activities that the team performs in the environment to accomplish the shared goals. Here we scope taskwork as the breakdown of the team mission into actions; essentially, the total workload to be completed. Taskwork, as we have explained briefly before, can greatly affect the context of what interactions will occur between humans and robots, [86, 87, 88, 36, 89, 90, 54]. The impact of this component can be detrimental or beneficial to teams depending on how well defined or structured it is. The lack of well-defined taskwork can set teams up for failure from a lack of preparedness if said team cannot work ad-hoc or is inflexible. Taskwork also creates workload not just physically, but also mentally on executing and or responsible individuals. For instance, one executing team member may continuously fail to complete a task, which may overload the team member that is responsible for the outcome of the task itself.

This formative component of teamwork has been well studied from the perspective of human factors, cognitive engineering, and cognitive psychology [36, 34, 33]. In particular, human factors tends to consider taskwork as a starting point of team design through top-down approaches. Taskwork breakdowns using methods such as hierarchical task analysis, and cognitive work analysis divide and categorize the overall workload [91]. Once the taskwork has been defined, it can be assigned through function allocation and concepts of operation to individuals or sub-teams [92]. Taskwork can also be designed by looking at the team's capabilities and the individual's skills it is composed of, a more bottom-up approach. While some teams may operate ad-hoc, others may be formulated and are highly dependent on the mission type and taskwork.

We propose the following sub-categories that better qualify distinct aspects of taskwork in Figure 2.4.

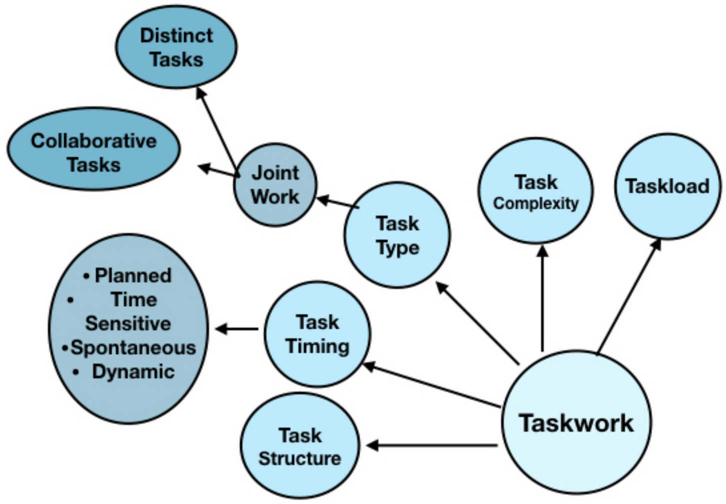


Figure 2.4: Mission Taskwork Concept Map.

Taskload defines the workload on an agent or across the team [33, 36, 34]. In particular, when dividing up the tasks, designers often look to create a balance of work across team members. This impacts which agent should be assigned work at what times and how much workload they will receive. The taskload received impacts other factors such as how much idle time, wait time, and working time agents have. Too much work may result in tired, stressed agents, while too little may bore agents.

Task Complexity refers to the simplicity or difficulty of a task. This is complexity impacted heavily by the capabilities of each agent to take on tasks of varying difficulty [93, 71]. Extremely complex tasks may require other agents' assistance, monitoring, or have high interdependencies with other actions that need to be completed alongside it. Other tasks might require multiple agents to coordinate, control, or otherwise assist each other. Some tasks may be simple enough for individuals depending on their abilities or motivation. Designers should clearly define the complexity of work and how it may impact each team member.

Task Type defines the classification of work, which can be distinctly individualistic (e.g. scavenger hunt across a city), closely collaborated (i.e. coworkers working in the same room), or jointly completed (i.e. two members solving one puzzle) [5, 94]. Soloist work required little to no interaction with other agents while completing the task at hand. Joint work, where two agents may work together and are tightly coupled [95], requires the most consideration of the different types of interaction that could occur between agents working so closely together. Joint work can be further broken down into collaborative tasks vs distinct tasks. Agents may play different types of supporting roles for joint tasks. Closely collaborated work may not require multiple agents to be hands-on for the same task but may require physical proximity or hand-off of work to others. These type of tasks suggests support for indirect interaction. Understanding the type of tasks agents will perform will help designers determine how proximal agents should be, how agents can support each other, and how closely collaborated or coordinated they may need to be.

Task Timing refers to the time sensitivity of taskwork [96, 71]. This indicates when the taskwork occurs and how long the task takes. Planned work and spontaneous work performed ad-hoc should be treated differently and their differences are taken into consideration. The timing of a task can greatly impact the success of the team as poorly timed taskwork can interrupt teamwork. Prioritizing tasks that are time sensitive can also disrupt teams if arranged poorly. Random situations, failures, blockages, and bottlenecks will have ramifications for task timing, and success in dealing with the consequences of pushed back timelines and delays.

Task Structure describes the arrangement and organization of taskwork [97, 5, 98]. Unstructured work may require teams to be more dynamic and flexible, while highly structured work may work better for teams that are less autonomous. The structure of taskwork also impacts how teams can formulate strategies to execute missions, i.e. military plays, bidding for taskwork, and ad-hoc teams. The structure is often highly influenced by the interdependencies of work, given that these dependencies can define the order of tasks or how parallelization of tasks can be done.

2.2.3 External Factors

External factors are aspects of the environment and system surrounding the team that cannot always be controlled or predicted. Different external factors can affect other teamwork components at varying levels of abstraction. The effects of these settings and conditions can also vary in degree or magnitude. A hierarchy of elements is organized in Figure 2.5.

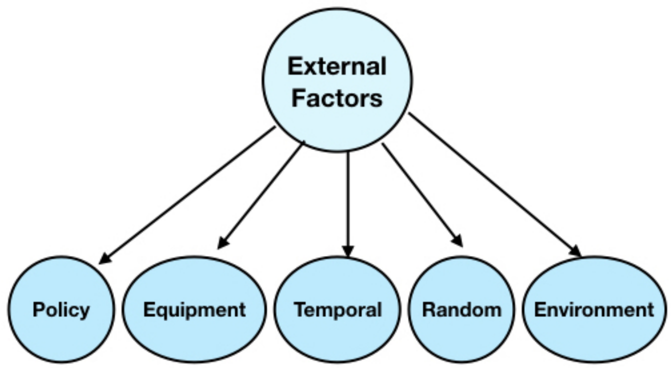


Figure 2.5: External Factors Concept Map.

While this is not a comprehensive list, we want to highlight the importance of considering factors that are directly external to the other teaming components we have mentioned. Recognizing the impact that some of these concepts we have assembled here may have to teaming will help human-robot teams be better prepared to deal with uncertainties that may occur. Additionally, considering the type of team or their ability to handle the accidental or unplanned may prove key to mission success.

External factors can also be controlled for depending on the mission circumstances. Therefore being aware of the various factors can help designers choose settings that have fewer variables, study the climate and terrain in full detail, and develop better tools that are more robust.

Public Policy Government and public policy may seem to be entirely unrelated or too abstracted away from day-to-day operations, but the reality is the power of protocols are able to restrict various components of teams [99]. This is not limited to the abilities of team members, how they are designed and tested, and the constraints on mission work that needs to be completed.

Equipment refers to the immediate gear, resources, and materials that a team utilizes as they work [71]. Physical objects may break, wear down, degrade, or go missing as a team is working. Additionally, teams may find arbitrary resources in their surrounding environment that may prove unexpectedly useful or pragmatic.

Temporal refers to unexpected time constraints or delays that may occur throughout systems or the mission [71]. These delays can hold back other agents from completing work, waste time, and delay mission completion. The stress of temporal changes can also cause harm to other resources, i.e. a scuba diver with not enough air to finish the dive.

Environment refers to both uncontrollable factors like bad weather (rain, wind, snow, visibility) as well as unexpected terrain (holes too large to cross, boulders too big to move around) [100]. If changes to the environment are unpredictable or unable to be observed, these pose real threats to the operability of robotic agents who may not be able to maneuver or scale these issues. For example, a human may not be able to cross a river that has risen many feet, and a rover may not be able to cross a field that is filled with boulders too tall to mount.

Random Factors refers to the factors that may occur due to random chance. This is a catchall to describe other accidental or contingent occurrences that are not encompassed in the other areas. Examples for this include unaccounted for individuals that may wander into the work environment.

2.2.4 Team Members

Human-robot teams are comprised of either humans or robot, which vary in a variety of factors such as skill, personality, training, and more. Though humans and robots have distinct traits that differentiate themselves from each other, they also share general qualities. Within humans and robots, human traits will differ from other humans teammates as no two humans are alike and robots will also differ from other types of robotic systems.

As team designers consider the individuals that will compose the team, the qualities of these individuals are important to keep in mind to determine team composition, what kinds of interaction may occur between members, and even what assignment of taskwork is given. There may be multiple agents who can accomplish the same task, but they may complete them differently, i.e. faster, more effectively, with less error. When you throw in the other factors we discuss in this framework, the range of performance may become even larger or smaller. There must be a consideration for the individual members because as we have discussed in the big picture relationship in the framework, they impact many other high-level teaming components. What kind of members are in the team and what sub-teams may form will change what kind of interaction will occur [50].

The qualities we highlight in Figure 2.6 for humans and robots are derived from cognitive psychology and robotics respectively. Here we make a point to describe not the differences but the shared qualities between humans and robots. By and large, many feel there is far more disparity today between the two than there are similarities. However, we make a case that this is due to the current capabilities of robots built today. As robotics hardware and software advance, we may find that robotics may be much more capable and more similar to humans that we see present-day. We first define their shared traits before discussing their differences.

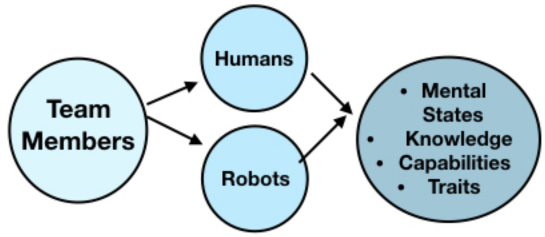


Figure 2.6: Team Members Concept Map.

General Qualities Shared Between Humans and Robots

Humans and robots share a surprising amount of commonalities, just at different levels of competence. For example, as a team humans and robots operate with certain shared mental frameworks. These frameworks allow for these individuals to have shared constructs of teamwork among themselves. Shared mental frameworks help facilitate shared knowledge, goals, understand, and situation awareness between each member.

We suggest that humans and robots can be thought of to have the following components, which inevitably factor into how they behave and interact. To understand the agents in the team thorough will create a reasonable expectation for individual and team performance. These individual components will categorize shared qualities between humans and robots. However, it is necessary to keep in mind the context of when a robot can utilize these traits and to what varying degree their capacity is for them.

Mental States differ when applied to humans and robots [101]. Both have mental states, although the states are of different degrees. A human mental state contains a person's beliefs, desires, and intentions, while robots' 'mental' or system states are subsidiary and contain their current status or surroundings. Humans have a much more complex mental state, whereas the mental state of a robot may be easily obtained and evaluated.

Knowledge we define as the acquirement of facts, information, and skills through experience or education, be it theoretical or practical [88]. Humans and robots both have knowledge. Humans learn and obtain knowledge as they grow from children to adults. While some robotic agents have simplistic software or interfaces, other robotic systems now have artificial intelligence through machine learning to allow them to 'learn'. This technology allows systems to declare what is contained in a picture, or how to maneuver around a space. Knowledge is also tied to the next shared quality, capabilities.

Capabilities we define as the ability to so some activity or process. Here though, we separate capabilities from knowledge in the actual ability to accomplish activities [71, 102, 60]. Both humans and robots have capabilities, although they differ in the specific type of capability. For example, both humans and rovers are capable of locomotion. However, humans have the ability to jump over rocks blocking their paths, while rovers must go around. Capabilities are impacted by the intelligence, problem-solving skills, and expertise of the individual. Understanding the shared capabilities and limitations of those for both humans and robots can help designers realize the full potential of each individual.

Traits we can define as the attributes, characters, and idiosyncrasies of an individual [103]. Humans typically have traits that are defined as personality, and these individualistic features are distinct to every person. While traits may seem odd to assume for robots, some robotic agents also have built-in personalities and quirks to them.

Limitations and Differences Between Humans and Robots

Though we have just described many crossovers within their respective abilities and traits, we have shown that humans are still more complex and have specific factors that today's robots do not yet have. The more complicated components of human personality we leave to other cognitive psychologists to describe in detail. In particular, human emotion and social factors can highly influence humans' roles, responsibility, and relationships with others in a team.

Recognizing the differences between future-day human-robot teaming could be and current-day limitations of robot team members are critical to deciding when to apply and consider these specific components. Given the technical limitations robots have from a hardware and software perspective, drawbacks in processing power and skills, like recognition or traversal, can impact the robot's performance, the rate of failures, and execution. The 'autonomy' of a robot and how it is defined can impact the robot's role and how other team members view it within teams. We present the concept of autonomy as a relative concept [50]. This implies the autonomy of a robotic system is not a built-in concept, but rather fluid dependent on different team compositions, mission taskwork, and circumstances in the environment. The variance in robotic control modes can overlap or change throughout the mission and is defined by manual control or command (by another human or robot). These differences and more offer a challenge to designers of when to activate which control mode.

2.2.5 Team Structure

Team Structure is a component of teaming that describes the organization and operations of a team. The design and assembly of the team configuration can impact an assortment of design considerations. As such, there are many ways to define a team formation or segment individuals within a team. This structure determines many crucial aspects of a team's working relationships and how they handle scenarios. The structure of a team answers questions like how prepared (or not) are they to handle randomness? Considerations for teams that behave randomly, or by a plan, can determine how disciplined or deliberate they operate? The structure is a good indicator of how the team behaves, and how designers can plan around that. We divide team structure into the roles agents may take on, the team size and ratio, the hierarchy of the team members, and organization within the team, Figure 2.7.

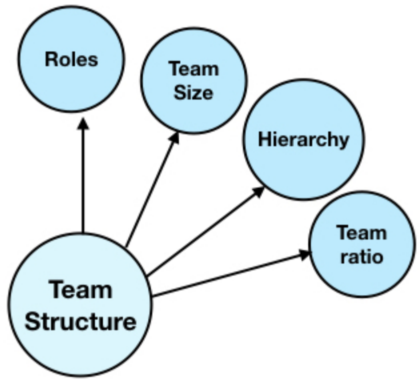


Figure 2.7: Team Structure breakdown.

Roles refers to the different types of responsibility and actions each team member may take on throughout the mission. Often a team member can take on a different role when required by changes in the taskwork, external factors, or even the team members they are working with. Some roles by Norman, Scholtz, and others have been defined as supervisor, peer, controlled, or bystander [57]. These roles factor into how responsibility is divided among members as well as relationships between members. Designers may need to determine an appropriate role (or roles) for each team member based on their capabilities, taskwork, and context.

Team Size is the number of team members on the team. Due to different member capabilities, finding the balance between too many or too few members is crucial for specific types of work. The team size may change depending on how many certain types of team members are needed or how many agents are needed to handle busy or slow moments. Team designers should evaluate the number of necessary team members, and acknowledge the scale of the mission to the size of the team. Overcrowding the team may cause issues with individuals stepping over each other while too few may overwork individual agents.

Hierarchy is the status or ranking of members relative to each other [5]. In some teams, each member may be equal to one another, but other teams may require or operate better under a chain of command. This hierarchy, whether it be ad-hoc or fixed, will impact the relations of the individuals and the success of the team in a given environment. Given a chain of command i.e. military rankings, then designers will need to understand how such a hierarchy can be best utilized to determine how members engage with each other or how communication across different levels may be impacted.

Team Ratio describes the proportion of different types of agents within a team. This can not only refer to the ratio of humans to robots, but also different roles, jobs, or skills sets. For example, the relationship between one human to many robots vs. equal of each vs. many humans to one robot will naturally require different design consideration. Certain roles like a captain or commander should belong to one individual (imagine two head of commands disagreeing with each other). Other roles or positions work better with many, such as an army or drones. Even jobs or skill sets can be considered; how many engineers does a maintenance mission require compared to strategy meeting? Designers should understand how to design team ratios to fostering effective inter-team relationships.

2.2.6 Team Interaction

Effective team interaction requires facilitating teaming concepts that occur as humans and robots work together, with respect to the work domain and limitations of the mission. Interaction, specifically within teams, has been studied through the lenses of many different focuses resulting in various perspectives of what interaction is defined as and what the key factors of interest are. Each literature area owns a specific view and analysis of human-robot team interaction.

In this section, we will describe many key interaction terms that have been well studied across theory, application, and experimentation. It is important for readers to note that while we present numerous interaction components here, we do not necessarily imply they represent a complete list of all interaction terminology. Instead, we hope to provide a catalog of terms that are most relevant and specific to human-robot teaming interaction today. Some of the aforementioned human-human teaming terms, which are key to human only teams, are not described here. However, we stress this is in part because human-robot teaming and robotic systems have not reached a point where those terms are yet relevant. The following conglomerate of concepts and terms in the teaming literature most pertinent to human-robot teaming as we know.

Common Ground/Shared Cognition: the concept of mutual belief having been reached by all parties about the meaning of the exchange [29, 76].

Shared Mental Model (SMM): is defined as organized representations of team and task knowledge that are shared and updated among teammates [77, 23, 78, 79, 80, 81, 104]. SMMs are considered a primary coordination mechanism because they enforce a shared understanding among teammates and ensure that the team has updated information about the critical components that are needed to facilitate joint behavior. SMMs include information about the task itself (*Task-SMM*) such as the environment, the required equipment, the operating procedures, strategies, etc. They also include information about the team (*Team-SMM*) such as the various roles, interdependencies, communication channels, and mental models of the individual teammates (their beliefs, goals, intentions, etc.). There is evidence that the team SMM directly impacts performance, whereas the task SMM has a more indirect effect [23]. A strategic consensus is a related concept to SMM. It is similar to SMMs, except with a focus on strategic priorities shared among team leaders. Whereas an SMM is shared among all team members, strategic consensus may only be available to managers or other individuals towards the top of the team hierarchy [24].

Interdependency: has been defined in various ways and acts as the central organizing principle between people [60]. Interdependency is not just the mutual dependence between agents in the joint activity, but the relationships as well. Johnson defines it as "the set of complementary relationships that two or more parties rely on to manage required (hard) or opportunistic (soft) dependencies in joint activity.".

Team Management and Leadership: is a process that involves one or more agents on a team that is responsible for executive functions, organization, and leadership within a team. Management tasks may involve defining goals, managing resources, assigning roles, and perform other influential tasks to ensure the functioning of the team. If undefined by the team designers, transition processes which involve the early stages of team formulation and planning can be carried out by leaders or those who manage the team throughout the mission [50]. Management also implies a responsibility for adapting the team when faced with changes and managing conflict by establishing behavioral norms which support other components of team effectiveness. Team leadership is specifically one or more agents on the team that is responsible for defining goals, managing resources and performing other tasks to ensure the success of the team.

Cooperation: is a high-level conceptual process by which team members work together to complete the same goal [50]. Cooperation can be thought of as both a guiding principle of design and a metric by which to measure team success. We propose cooperation as a concept that is internal to agent processes but does not automatically imply any direct trust or appreciation between team members.

Directability: refers to "one's ability to direct the behavior of others and complementarily be directed by others. Directability includes explicit commands such as task allocation and role assignment as well as subtler influences, such as providing guidance or suggestions or even providing salient information that is anticipated to alter behavior, such as a warning" [60].

Trust: is one concept that has been defined multiple times from various source. Mayer suggests it is the belief that team members will fulfill their part of the task without the need to monitor or intervene [22]. From others, we can presume trust to be a voluntary and purposeful delegation of goal-oriented behavior derived from a dynamically learned expectation of the trustee's consistency of performance and compliance with norms and laws, conditioned on a dispositional evaluation of the trustee's internalization and adoption of the trustor's goals, as well as trust pre-dispositions, the current situational context, and the potential costs and risks of failure [72, 73, 74]. Mutual trust, in particular, is a shared understanding that teammates will perform their part of the task in a timely and appropriate manner. Fong et. al. states trust may influence an agent's reliance on complex, imperfect automation in dynamic environments that require the human to adapt to unexpected circumstances [71]. More specifically, humans must trust that a robotic teammate will protect and welfare of the team and its individual members. The level of trust among team members is also critical for high-risk situations [105]. Trust also affects the willingness of human team members to accept or follow robot suggestions or assistance [70].

Fluency: has been defined in many ways. One definition is collaborative fluency, essentially the quality of the interaction in joint activity [75]. Fluency is closely tied to trust as well as situation awareness.

Reliance: the act of dependency and assurance or believe in the act of others to perform based on one's expectations[106]. Lee suggests that reliance (an attitude) is impacted by trust (a belief).

Prediction: the ability to predict other actions that also encompasses the use of models and mechanisms in order to do so [60]. Being able to predict others as well as being aware of how to make your own actions predictable to others is important to teamwork patterns and creating efficiency in team performance [60].

Closed Loop Communication: refers to more than communicating a thought, action, or intent from one end to another. The specification of the closed loop implies that not only should the listener acknowledge the message, but the speaker should also confirm that the message was correctly interpreted [107].

Physical and State Signaling: refers to the active signaling of an agent's state to another, through whatever means they are capable of [108].

Flexibility the ability of the team to deal with unpredictable and dynamic circumstances, allowing them to work synergistically with each other, systems, and environment [109]. This is a factor in determining team success through their ability to continue or not past various types of interruptions to their taskwork or overall mission, typically in the short run. Improvements in flexibility result in less brittle and more robust teams. Strategy Flexibility is akin to flexibility (physically) but refers specifically to the overall strategy and operations of the team.

Adaptability: is the ability to recognizing changes or issues as they arise and re-adjusting individual or team behavior to compensate. The issues can be internal to the team (e.g., conflict, failure, etc.) or external (e.g., environmental change, contingency, etc.) [65]. This is deeply enmeshed with the concept of flexibility but is distinct due to its long-term implications. This requires all of the components described above, including maintaining an SMM to anticipate how the new changes will affect the team/task, monitoring teammates to determine if adaptation is needed, backup behavior to assist and accommodate as necessary, and leadership to initiate and manage these adaptations.

Joint Attention: here refers to the following of cues (visual or otherwise) by two or more team members [110]. Much of the focus on visual cues come from gaze towards objects in a shared environment. However if extended to HRI, joint attention could also apply to remote objects, information, and taskwork. Cues from the initiator then offer clues to their intentions and goals in order to facilitate comprehension [110].

Awareness: is the concept of being cognizant and knowledgeable of the physical world. It can be broken down into the category of automation that a human is working with [88]. For HRI, that is human-robot, human awareness can be considered " the understanding that the humans have of the locations, identities, activities, status, and surroundings of the robots" [88]. For robots, other research has defined situational awareness, which changes depending on the task and the ability of the robot [84]. At a broader level, it is the knowledge of what is going on around oneself [111]. Different perspectives also impact the level of situational awareness one can have.

Observability: can be thought of as "making pertinent aspects of one's status, as well as one's knowledge of the team, task, and environment observable to others... Observability also involves the ability to observe and interpret pertinent signals" [60].

Recognition: the identification of an act, item, process from prior engagement into a preconceived thought or understanding. Recognition measures also include classification and familiarity with that are being observed [100].

Comprehension: the ability to understand the environment, situation, or meaning of something [88].

Interpretation: is a process by which "meaning is extracted from ambiguous information to construct a mental representation" [69]. This is the act of conceptualizing one's own perceptions.

Judgment: the ability to form a decision based on the knowledge and perceptions that an agent may have. Judgment is also a process by which an individual will consider and "evaluate evidence to estimate a different likelihood of different outcomes" [69]. Judgment tends to be a more personal quality, formed by the skills acquired by individuals over time.

Reasoning: the cognitive process of forming inferences or understandings based on the information available [69]. Reasoning is the formulation of explanations or

Decision Making: the ability to form or choose between different options [112]. It is how people can choose one of our several options with a focus on how they select or avoid specific options (that may have risk) [69]. This factor involves a certain degree of conflict and stress when difficult decisions are involved.

Planning: the ability to form a strategy or arrange activity prior to or during the mission [100].

Climate: represents a team-level state which has to do with the attitudes and expectations by which the team operates. Climate has multiple dimensions and has been broken up into safety, service, and justice climate. *Safety climate* refers to team attitudes about safety procedures, and perceptions of safety climate have been associated with the rate of accidents and injury [18]. *Service climate* refers to organization-level concerns about customer satisfaction and is associated with perceptions of service quality [19]. Finally, *justice climate* refers to a team-level sentiment about how the team is treated, and it is associated with performance, commitment, and absenteeism [20].

Team Confidence: refers to beliefs about competence as it relates to a specific task (efficacy) or more general confidence (potency) shared by all team members [7]. These states have been correlated with improved team performance [16]. A related concept is *Team Empowerment* which is split into two main categories - structural and psychological. Structural empowerment deals with how the structure of the team can lead to effects on performance and satisfaction, whereas psychological empowerment deals with either an individual or collective sense of authority in terms of controlling outcomes. Psychological empowerment is associated with improvements in performance and customer satisfaction [17].

Cohesion: is a team-level state which emerges somewhat late in a group's operation and broadly deals with the unity or bond that is formed in the team. Cohesion has been decomposed into three component categories of interpersonal attraction, task commitment, and group pride, and has been strongly associated with effective team performance [21].

2.3 Human-Robot Teaming Interaction as a Model

Many of the terms that people use to study and define interaction are important, there are limitations to how well using a subset of these terms can help one can understand team interaction fully. Generally, HRI does not have a strong classification system for how these terms relate to or impact one another. To understand how these terms relate and impact on another, we must map them to an identifiable construct of interaction.

Figure 2.8: Scholtz's HRI model of goals, intent, actions, perception, and evaluation.

One way to better grasp the complexities of HRI is to represent interaction as a model. Scholtz defined a model of HRI in teams which define an individual's internal movement through the development of goals, intentions, actions, perceptions, and evaluations, see Figure 2.8 [57]. Besides being an HRI model that many have built studies and experiments around, this model creates a way of looking at HRI that is independently defined by many of the interaction terminologies we defined previously. These cognitive concepts are defined as a cyclical process inherent to each team member. Thus this model allows us to understand the internal processes an agent may have while interacting with others. The progression through the cycle of the Scholtz components is also volatile depending on other factors like team structure (roles of agents) or team members (capabilities, skills). This model will serve as the stepping stone for much the Interaction Framework and Methodology we will present later on.

How then, might we tie together the model of interaction with the numerous HRI terms we've gathered from past research? We need to organize the interaction terminology onto this model by fitting each term 'into the bucket' that fits most from the conceptual model. To do so we must first expand the Scholtz components and resulting model. This expansion lets us define and expand this HRI model into each component as well as the process of moving between them. This addition of transitional states between the original five states opens up the model to more details. While the original states are clearly separate and distinct, there are more factors that go into the process of an agent progressing from one to another. We define these five states and five transitions below:

Goals represent what the team and individuals want to accomplish.

Goals to Intent represents the formation of intentions from goals.

Intent is how individuals intent to satisfy the goal.

Intent to Action is the formation of actions from intentions.

Action is the execution of an agent's intention.

Action to Perception is the perception of actions by themselves or others.

Perception is an agent's assessment of actions and ongoing surroundings.

Perception to Evaluation the formation and understanding of an agent's observations.

Evaluation is the comparison of an agent's system state to their intention.

Evaluation to Goals the potential use of information to update and change goals.

We present an Interaction Model, which provides an understanding of high-level relationships between the HRI concepts and components in Figure 2.9. Based on the definitions and our understandings of the interaction components in Section 3.6, we allocate the components to each of the Scholtz's main and transitional areas. Additionally, the placement of HRI terms on this model allows us to visually understand how each concept may be analogous or complementary to another. Here, the concepts that lie in the transition states are not necessarily unique to either but are rather a conjunction of both.

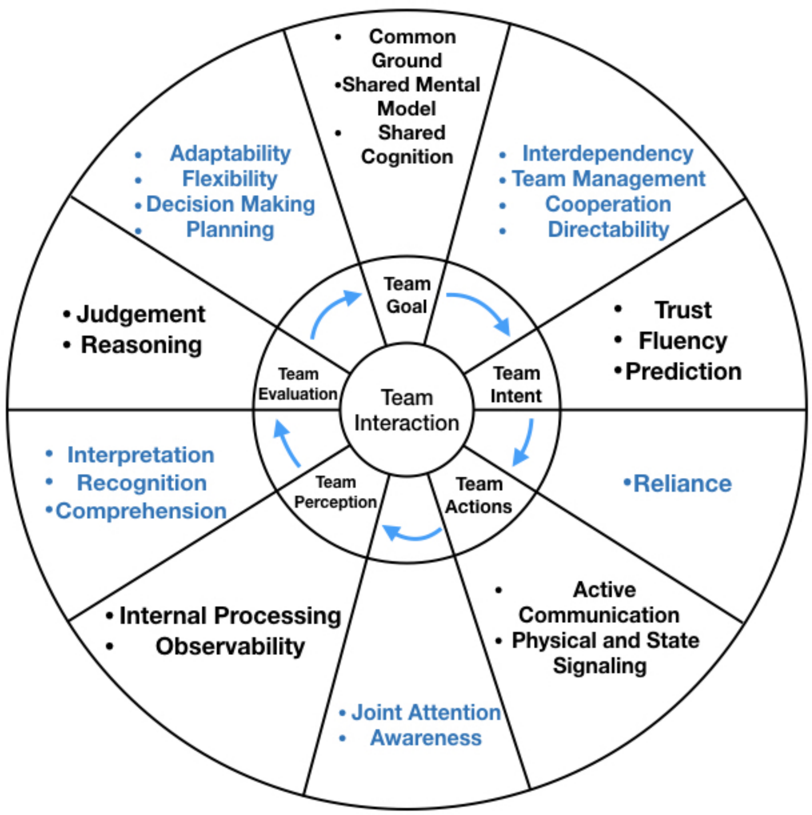


Figure 2.9: Team Interaction Model

2.3.1 Interaction is Central to Human-Robot Teaming

Human-robot teaming is defined by the meta-interaction of the components we've identified before: taskwork, team structure, team members, and external factors. The prior general framework describes the five main components of teaming and represents each aspect of teaming as equals, see Figure 2.2. However, it is easy to see within our findings and the literature that interaction is seemingly more complex than the other components. Compared to the other components of teamwork identified, components that define team interaction remains the most heavily debated and unregulated. Some research debate the other components (taskwork, environmental factors, team structure, and team members) are the precursor to what we should expect from interaction within a team. Due to the fact that many of these components can be predicted or defined to some degree prior to the mission performance, these factors can be considered inputs that really determine what kind of interaction will occur between the team members.

Thus, one perspective of this general framework is to focus the other components around Interaction. Our description of teamwork (having a shared goal, working together, working effectively) remains in this new interaction focused framework. Here while we show that the other teaming components remain important to the Framework, Interaction takes the center stage. We re-frame the holistic view of human-robot teaming to focus on Interaction as the key factor to consider among team members, team structure, taskwork, and external factors. The shifted framework of human-robot teaming to centralized interaction as the leading component of teaming is shown in Figure 2.10.

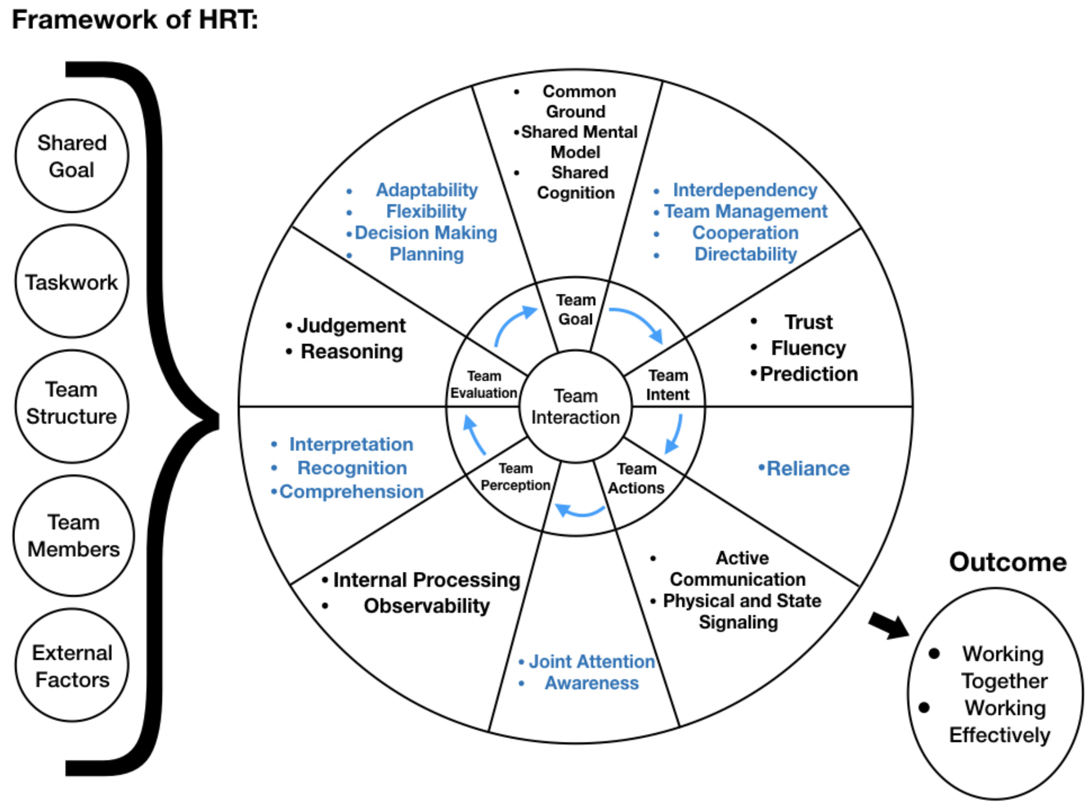


Figure 2.10: Interaction as the focal point of Human-Robot Teaming.

One should recognize that not all the concepts listed in this section are of equal value to every human-robot team. Indeed while many of these components of interaction are core aspects of team interaction, not all teams need to consider every component. This interaction-centric framework provides a broad organization of key components to understand what affects teaming and teamwork. More importantly, understanding the different levers that can affect the team and their impact provides an early-in-design method of facilitating effective teaming. Every team and mission will be different and require understanding different proportions of various parts of this framework.

This framework is most helpful for understanding the overall impact that HRI factors have on each other and agent relationships. Team designers cannot only confirm whether their areas of interest fall under the larger umbrella area of focus, but also what concepts might be missing from their design. While the frameworks describe the design considerations in detail, one specific point of interest may be how it applies to existing design methodologies. Here we will show how this framework spans across multiple design methodologies that highlight distinct aspects of HRI.

2.3.2 Applying HRI Methodologies to the Interaction Framework

Earlier we mentioned how many HRI researchers choose specific HRI terms of interest to study. This selective process is also seen in theoretical and applied methodologies for HRI. A common practice for evaluating HRI in teams is through a specific methodology. Specific here means that some of these methodologies are defined to evaluate a discrete set of interaction factors through its design. While these methodologies are not extensive to evaluating human-robot teams holistically, they do provide value to designers by deeply investigating specific team characteristics. Some "methodologies" are really evaluating systems and are not suitable for the formative design process [113, 114, 115, 116, 117]. Much rarer are methodologies that target formative design and aid designers in assessing the concepts in Figure 2.10.

To show how and which terms these methodologies cover, we will describe several methods and the subset of terms within this HRI framework they each cover. Here, we demonstrate the wide breadth of HRI terms and how a single methodology, while still useful in its applications, does not necessarily span the entirety of HRI terminology. We overlay the components that each methodology focus on in Figure 2.10 to show how various methodologies do not cover all the components of teaming.

Coactive Design

Johnson's methodology highlights interdependency its core which shapes autonomy and defines relationships between team members. He defined three main elements to determine interdependencies: observability (red), predictability (yellow), and directability (orange) [60]. These cover a spectrum of the HRI model, represented in Figure 2.9, these bars cover the area of which their impact is expected, see Figure 2.11.

As we define these terms and their place on this figure, we bring attention to the concept of full-circle or closed-loop components. Specifically, the definitions of observability, predictability, and directability imply not just the direct action from the outgoing team member, but also the recognition from the receiving individual. OPD offers a way for designers to identify which teamwork requirements may need more support.

Observable (red) elements can be applied from building intention to execute actions to a member's perception of their own and others' actions. Observability means "making pertinent aspects of one's status, as well as one's knowledge of the team, task, and environment observable to others" [60]. This implies coverage from Intent through Evaluation, as making something observable is only a part of observability. Johnson notes that interdependence is about complementary relations, and thus observability also involves the ability to observe and interpret pertinent signals [60].

Predictability (yellow) is how agents can perceive another team member's intentions and perceptions. "Predictability means one's actions should be predictable enough that others can reasonably rely on them when considering their own actions" [60]. We span predictability from Intent up until Evaluation as predictability is involved when agents anticipate other member's intents and actions, but also when agents process whether their predictions were correct or not. Therefore predictability captures both matching expectation and inference from actions.

Directability (orange) is largely expected when forming goals to intentions, particularly in aligning shared goals and management of team intentions. "Directability means one's ability to direct the behavior of others and complementarily be directed by others" [60]. In our framework, directability covers the transition from Evaluation to Goal through Intent to Action. Directability here can span direct manual control to task and role assignment to subtle influences such as guidance or suggestions.

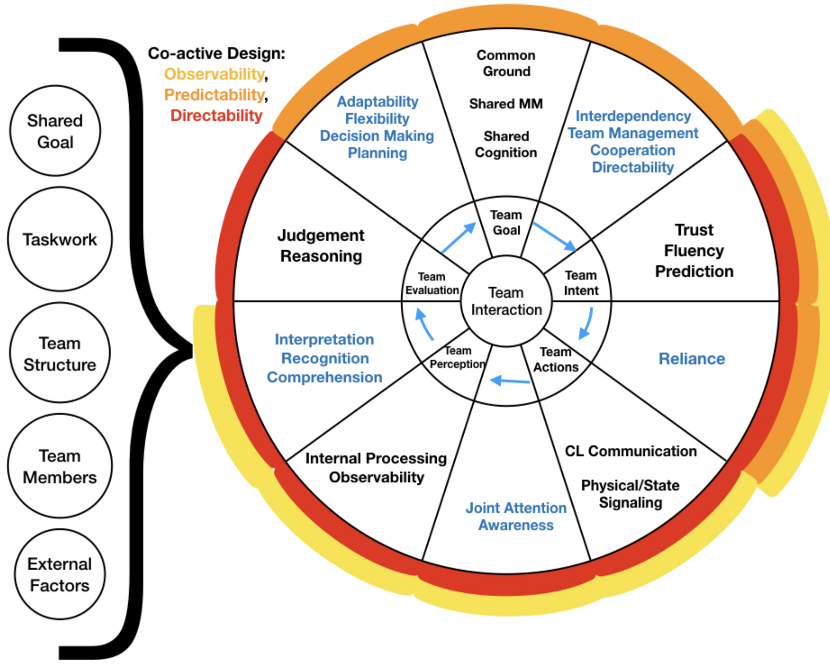


Figure 2.11: Fitting coactive design onto the HRT Framework

Objective Work Allocation

Within human factors, one methodology that has been well studied is the application of work allocation to human-robot teams. Work allocation is the team design of assigning different tasks to different agents [36, 33]. Designers often start by identifying the taskwork required for the team to complete its mission, either through Hierarchical Task Analysis or Work Domain Analysis [44, 47]. These analyses may identify a single sequence of action, or they may identify feasible groups of actions as important to the mission or team flexibility in how the mission goals may be achieved, and how variable the task environment is.

In particular, this methodology is based on computational simulations of human-robot and human-automation teams. The goal of this method is to provide informed design decisions based on the team composition, allocation of work, and the interaction protocols [118, 52, 51]. The main advantage of computational simulation here is, if human-robot interaction is modeled properly, the ability to capture emergent behavior in the team to aid designers plan more effective teams.

This methodology is able to perform analysis of constraints inherent to the taskwork, environment, and agents, explicitly capture human-robot teamwork and computationally evaluate the impact of the interaction between the modeled components of teaming. The methodology is split into three phases: pre-simulation analysis (breaking down the work and agent attributes), simulation, (dynamically observe a human-robot team's performance over time), and post-simulation analysis (analysis of the simulation output).

Due to the origins of research that spawned this methodology, the focus of this method is taskwork definition and breakdown. The detail of which the mission taskwork is analyzed and allocated is key to this method. Among taskwork, external factors, environment, and action interaction are also important factors to this method. These components are noted in blue in Figure 2.12.

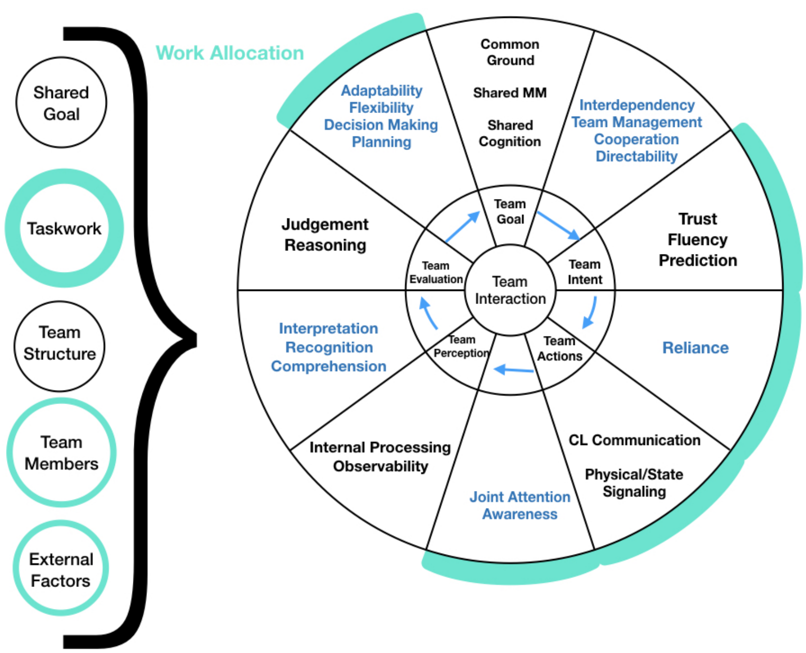


Figure 2.12: Fitting objective work allocation design onto the HRT Framework

Interaction Design Patterns

[119, 120] Interaction design patterns provide designers with a specific format to both acquire and disseminate design knowledge with respect to the specific context surrounding a problem in teams [120]. Specifically, a design pattern describes a repetitive problem that is reoccurring as well as its solution. One methodology that uses interaction design patterns is situated in Cognitive Engineering (sCE), which is applied to investigate human-robot collaboration. This iterative methodology is a "human-centered development process, aiming at an incremental development of advanced technology" [120].

The foundation of this methodology, in particular, is based on earlier views on Cognitive Engineering and focuses on three concepts: foundational knowledge, specification, and evaluation. Foundational knowledge identifies the contexts of the environment, specification refers to the refinement of requirements, and evaluation validates and tests the requirements. Within specification, there exist two levels that determine the communication level (interaction design) and task level (user requirements and design rationale). The communication level specifically utilizes design patterns to define the "how" that is derived from first investigating the task level.

The sCE utilizes design patterns to solve interaction problems. Interaction Patterns more specifically describe an approach to identify and solve problems that arise in the taskwork, context, environment, and communication between the teammates. We can map these components on the HRT framework, highlighted in green in Figure 2.13, with a focus on the taskwork and environmental constraints.

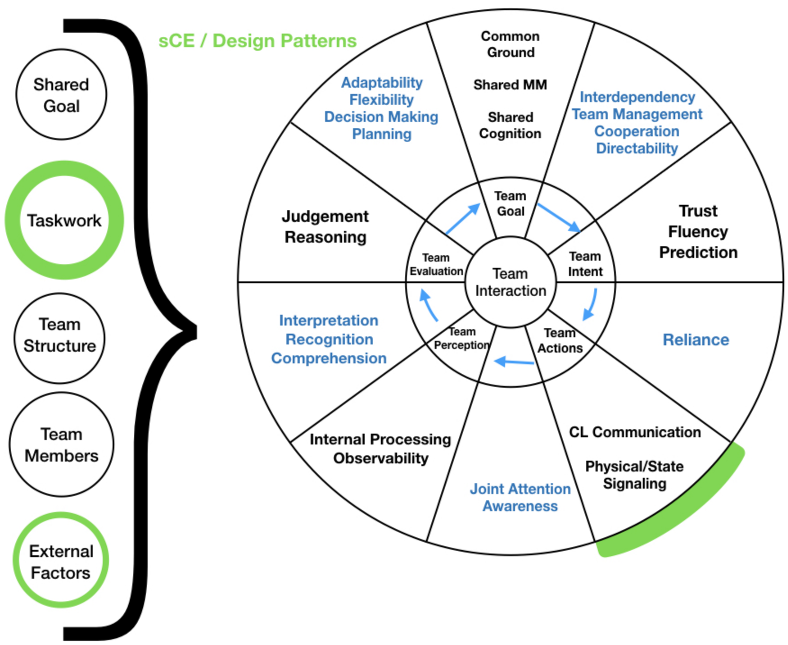


Figure 2.13: Fitting interaction design patterns onto the HRT Framework

IMOI Framework

The Input-Mediator-Output-Input (IMOI) framework (introduced in section 2.1) provides a very general overview of the core requirements of effective teaming. In Fig. 2.14, we have highlighted in purple the components of the IMOI framework that corresponds to our HRT framework. In particular, these include the Input to the team and the Output (outcome measures), as well as the process and emergent states. Among the many team processes that have been identified in the literature (see [13]), the key ones that align with our framework include decision making, planning, cooperation, team management, active communication, and joint attention. The key emergent states that align with our framework include SMMs, trust, reliance, observability, interpretation, and reasoning. Overall, the IMOI framework generally much of the inputs as well as the evaluation to action area of our HRT framework. However, our framework makes additional theoretical commitments that are not present in IMOI, namely the Scholtz model. On the other hand, IMOI lists a vast number of mediators that link Input to Output, but it does not otherwise describe how these mediators work. This is problematic, especially for human-agent teaming, because it is not clear which attributes of teaming are necessary for which kinds of outcomes. Our HRT framework is a step in the direction of building upon IMOI in order to make it useful not only for understanding but also supporting the types of interactions necessary for effective teaming.

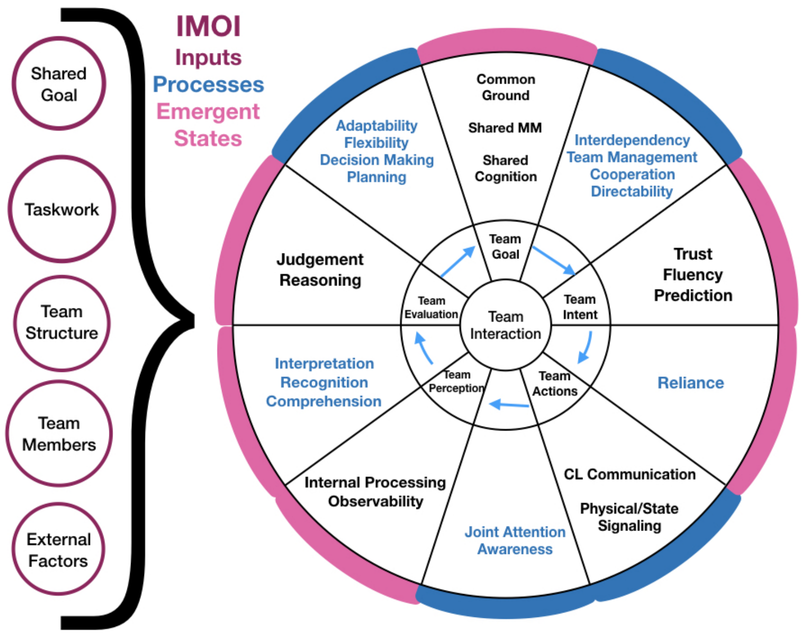


Figure 2.14: Fitting the IMOI method design onto the HRT Framework

2.4 A Methodology for Studying Human-Robot Teams and Interactions

Although we demonstrated the functionality of different methodologies for investigating HRI in human-robot teams, it remains difficult to objectively compare these different methodologies. This is in part due to the lack of consensus for human-robot teaming as a whole as some point out it is not clear what a foundation of HRI methods would look like [56]. While the aforementioned methodologies allow for the in-depth evaluation of a specific set of HRI metrics and concepts, there is no singular method for investigating every aspect of teaming that would be relevant for the team in question.

We offer a new methodology here not as a catch-all to evaluate human-robot teaming, but as a way to holistically consider the five components of teaming in detail to more effectively investigate human-robot interaction components of interest. This methodology involves the merging of the Teaming Framework from Section 3 with the Interaction Model of Section 4 in order to better understand which HRI components may be important to success. We apply our method to various different types of teams through four main case studies. The case studies range from simple two-man human teams to more complex multi-agent teams taken from existing human-robot teams in HRI literature. We hope to show with the varied case studies how representation in this method will not only confirm key interaction factors found in the original case studies but also bring to light interaction components that may have been forgotten or overlooked. Additionally, these case studies will demonstrate the flexibility of this framework to various types of teams, as well as highlight the different types of dynamic interaction that can arise.

More specifically, this method takes the four factors of teaming (team members, team structure, taskwork, and external factors) and applies them to the Interaction Model to generate a teaming model. While earlier we pointed out how inter-related the five teaming components are, we stress an order of dependency by which to define each component 2.15. Each component can be defined through the main characteristics we detailed in Section 3. External Factors and Team Members can be defined somewhat independently, and therefore first. These two components both may impact the Taskwork and Team Structure, which can be defined in parallel as they can influence one another. Finally, the definitions of these two can be combined together and applied to the Interaction Model (Section 4). This model will be expanded into a dynamic representation of actionable tasks as well as compared to the Interaction Framework.

This methodology can, therefore, be distributed into four main steps: 1) understanding and defining the mission work, external factors, team structure, and team members, 2) creating an interaction model of teaming using factors from taskwork and team structure: proximity, motivation, and interdependencies, 3) adjusting the interaction components for

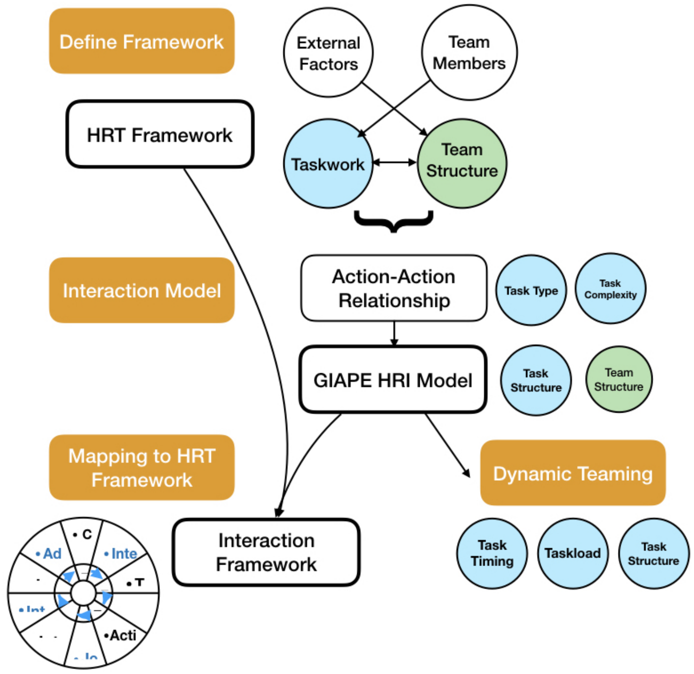


Figure 2.15: Visual flow of conceptual methodology.

each action in the taskwork sequence to capture dynamic teaming changes and weights, and 4) mapping the interaction model recommended components to the interaction framework(Section 4) to showcase integral interaction components specific to the team. This methodology for applying the framework of interaction is in more detail below:

1. Defining Teaming Factors. Here we define the mission and other 4 main factors (team members, team structure, taskwork, and external factors) clearly. These definitions should follow the breakdown of each concept from Section 3.

    - First, define External Factors (random, environment, equipment) and Team Members (capabilities, traits).

    - Breakdown taskwork for the mission into specific actions and determine subcomponents (taskload, complexity, type, timing, structure).

    - Breakdown team structure for the mission (roles, size, hierarchy, ratio)

2. Create Teaming and Interaction Model using Scholtz's HRI model (Section 4).

(a) Generate the Action Relations structure based on the breakdown of taskwork from step 1. This sub-step formally categorizes the task type, complexity, and structure.

- Task Type: we define three main categories of task type as joint work(tightly collaborated, close proximity actions), closely coordinated (near proximity actions done in sequence or with coordination), and remote work (actions done separately from one another).
- Task Complexity: we define through the requirements or motivation of the action, specifically those that may require others assistance. Motivation determines how the action is perceived or controlled by other team members, which falls into monitoring, commanding, and self-driving (internal drive).

(b) Derive Interaction Model (GIAPE) from combining the learnings of Step 1 and Step2a. The impact on the other GIAPE elements depends largely on the team structure (roles, hierarchy) and task structure.

- Roles and Hierarchy: the role that each team member can take on throughout the course of the mission. This defines the team hierarchy as well. The role of an agent impacts which GIAPE elements are able to be self-influenced, commanded or monitored by others.
- Task Structure: is impacted by the roles and hierarchy. Here we refer to how this structure impacts dependencies of the GIAPE elements of one another.

3. Create dynamic action model of teaming. This step formally categorizes the other aspects of Taskwork we could not capture statically, task structure, timing, and taskload.

(a) Breakdown each agent's action in time sequence relative to each other.

(b) Draw relationships between actions, translating the model to this dynamic sequence.

- Task Structure: here we specify as the interdependencies of actions to other actions. Interdependency here refers specifically to how much reliance tasks have on other tasks to be finished, started, or done in conjunction.
- Task Timing: is shown as each action is laid out over time. Here we can see which actions can be done in parallel or linearly.
- Taskload: is the workload on each agent. This is visually represented as the allocation of taskwork to agents to shown over time.

4. Map and compare the interaction framework overview of the generated model of teaming. Here we finally compare the Interaction model we generated with the Interaction Framework (Section 4). Any crossover between these will show meaningful interaction components (Section 3).

2.4.1 Applied Methodology

To begin, we will describe the methodology alongside a demonstration with a short, simplistic case study: two humans moving a table. Afterward, we will show several more complex teaming examples to demonstrate the framework in more detail and breadth of team dynamics.

Step 1: Defining Teaming Factors (members, structure, taskwork, and external factors)

In this initial step, the clear (or as clear as possible) characterization of the other four factors of teaming (team members, team structure, taskwork, and external factors) in accordance to the branching terminology breakdown is fundamental. Any external factors at play should be defined as thoroughly as possible. Team members involved in this short mission should be clearly stated along with their capabilities/traits. The team structure should be well defined (roles, team size, hierarchy, and ratio). For taskwork specifically, we will break down the mission into its specific actions, as well as the taskwork factors. Here to better describe the methodology, taskwork will be relatively high level and simplistic and we will forgo complex external factors.

External Factors : the setting is a plain living room. Let's assume in the simple example there are no extreme external factors (pets running around, table breaks, etc).

Team Members : Two human adults from a moving company. Each has its own internal process behind interaction, and therefore capable of understanding and defining their individual goals, intentions, actions, perceptions, and evaluations. Both are able-bodied and for this example, similar in physique.

Team Structure : The humans' roles are both peers and they are considered equals. They do not operate with a hierarchy amongst each other, and the ratio of the team is 1:1.

Taskwork : To lift a table together and set it down. The overall mission this task belongs to could be a moving company helping a customer move their belongings around. For this simple case, we look at the specific set of actions involved with the task: moving a table.

1. Hold table edge opposite each other.
2. Lift edge of the table together.
3. Walk together to a new area while balancing table.
4. Set down the table.

Task Type: The entirety of the task is a joint activity, within close proximity to one another. No person can lift the table alone.

Task Complexity: The overall task is relatively straightforward. The complexity lies in the reliance on each other to monitor and occasionally command one another if they need to move around a blind spot.

Step 2: Create a Teaming and Interaction Model

Now that we've laid out the details of the teaming factors, we can assemble the team and interaction model. Finding the interesting components of interaction relies not only on the definition of external factors, members, and team structure but a deeper evaluation of relationships that can form agent to agent. The actions of an individual throughout the mission can be a driving force behind resulting interactions or rapport. Therefore before we can derive a model of the teaming, we must first thorough examination of how individuals may decide on or perform actions. To model the effect actions have on the individuals' internal processes, we utilize the constraints by which we can measure the types of relationships each agent can have to another through the taskwork components. These components at this stage we limit to task type, complexity, and structure.

To begin, we represent the actions of agents as circles. The relationships between these actions are drawn through different lines that differ in types, colors, and weights with respect to task type, complexity, and structure, see Figure **??**.

- Line type: represents task type and has three different lines being joint (solid), close (dashed), and remote (dotted). As task type here also relates to spatial proximity, joint actions will share overlapping areas, close actions will share overlapping areas with a dashed outline, and remote actions share no overlapping areas.

- Line color: represents task complexity and the requirement of other agents to assist with tasks. We classify three types as monitoring (green), commanding (red), or self-driving (blue). Actions that have multiple relations to another will have multiple arrows to show this.

- Line weight: represents task structure and dependencies. We show this in here legend but will not implement this until we view dynamic visualizations of interaction.

For our mission, moving the table, we see the following representation where 1) Line type: Solid as the entire task is a joint activity and 2) Line color: The complexity lies in the reliance to monitor and command each other if needed as well as self-monitoring, see Figure 2.16. We know spatially the proximity of two humans lifting the table will be close and joint. The drive of actions effect on one another is some form of monitoring or command.

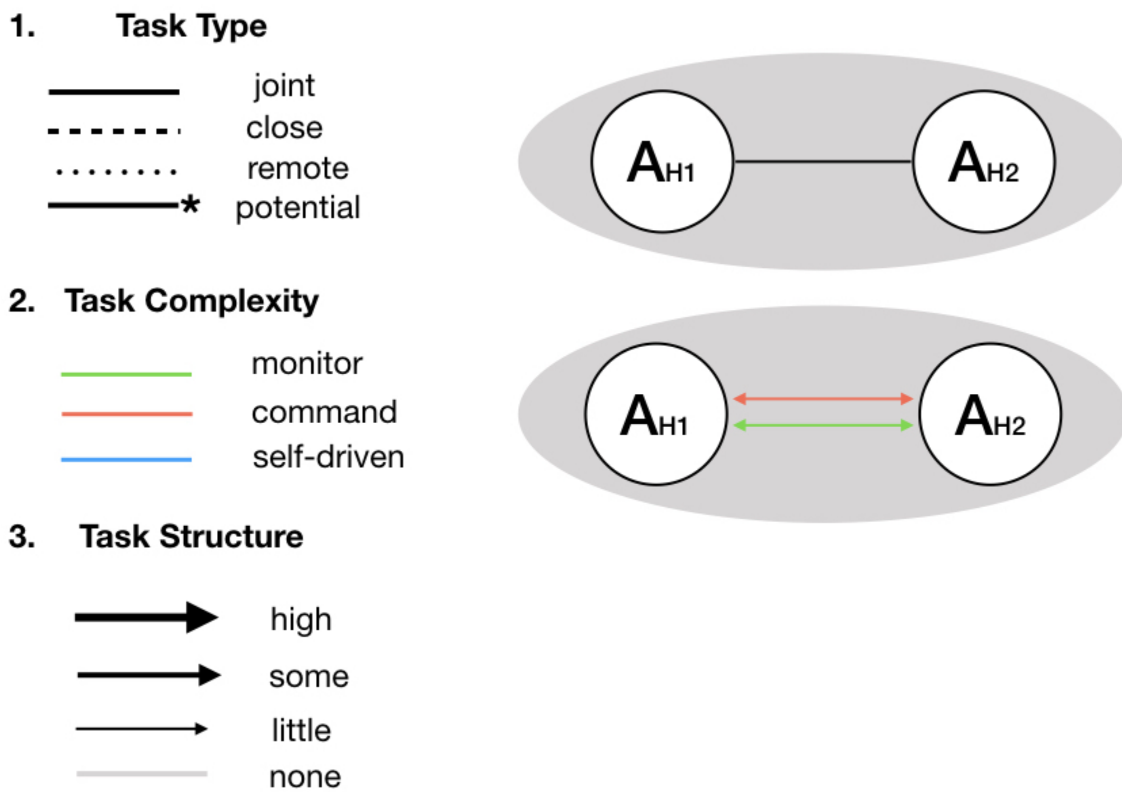


Figure 2.16: An example of the action relations between two human agents with a legend (left).

Now that we've clearly defined the action to action relationships, we can apply these relationships to the Scholtz HRI model. In this step, we show the action component expanded as we've just defined it. The relationships between each of the other components (goals, intent, perception, and evaluation) follow from the taskwork characterizations. In addition, we can now apply the definition of task structure, more specifically, roles, to this model. Role definition was a central theme behind the organization and bridging between these HRI model components. For example, a simple robotic agent would not be able to set its own goals, whereas a human operator would be able to set the overall mission goals for all agents involved. Team structure here also impacts the hierarchy or chain of command, or rather how information flows through agent to agent.

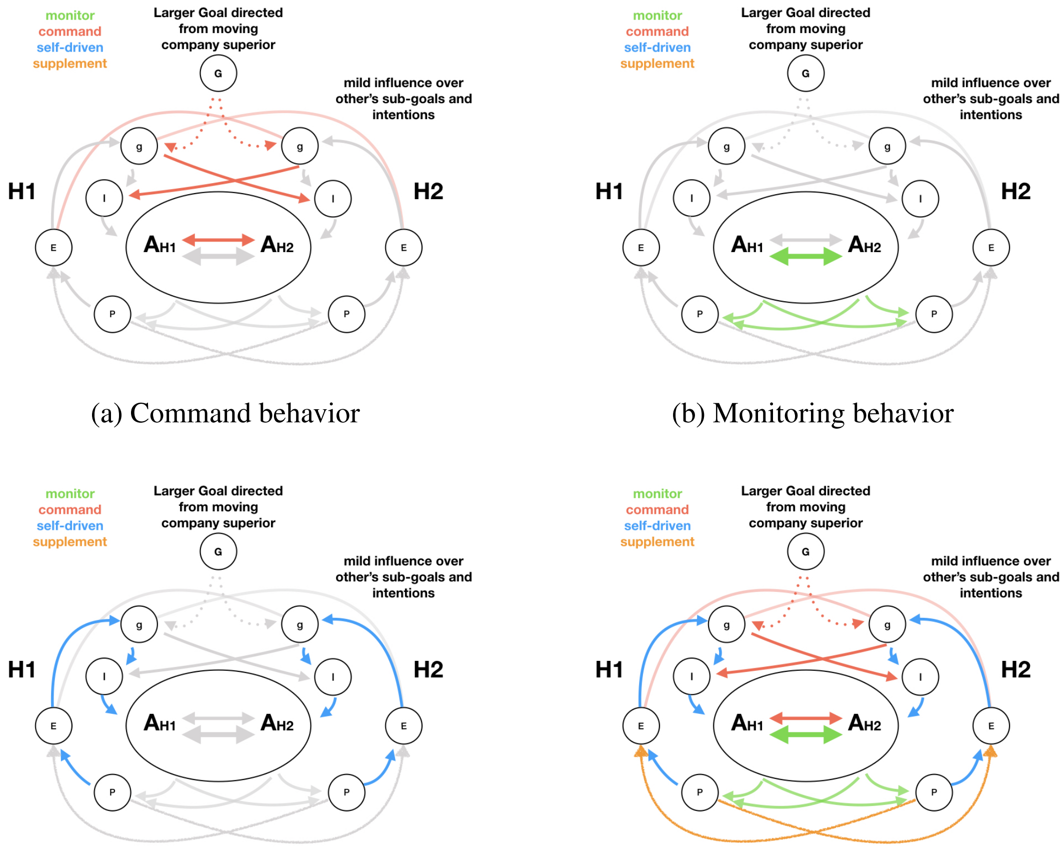
The representation of the Interaction Mode through GIAPE elements is shown in a series of four steps in Figure 2.17d. Here we see a mission overview of how these tasks and team structure impact the transitions from goals, intentions, perceptions, and evaluations.

- Starting with the Goal: Here “G” refers to the larger mission goal, while “g” refers to the sub-goal inherent to each individual. We abstracted away earlier the idea of the goal of moving a table to the company superior. Thus each individual human has their own subgoals associated directly from the larger goal G. These ‘g’ goals have some degree of influence over each other’s intentions to execute this goal. Additionally, the evaluations they make are capable of impacting each other’s short term goals (i.e. commanding one to avoid bumping into the wall/other objects). This command is shown in red, see Figure 2.17a.

- Next we move through Actions and Perceptions. As they work through the actions to complete the overall mission, each agent will monitor one another, directly impacts their perception of the situation. This monitoring is shown in green, see Figure 2.17b.

- Both agents are humans, with their own self-autonomy to process the full loop. While they are not allowed to change the large ‘G’, they are able to impact their own goals ‘g’, intentions, actions, perceptions, and evaluations. The feedback loop from their evaluations helps each agent determine their next subgoal ‘g’ over time. This process is shown in blue (self-driven), see Figure 2.17c.

- Together, these components combined form the Interaction Model that is derived from the definition of the teaming components identified in Step 1, see Figure 2.17d.



(a) Command behavior

(b) Monitoring behavior

(c) Self motivation

(d) All

Figure 2.17: Breakdown of teamwork model into command, monitoring, self motivation, and a combination.

Step 3: Dynamic Action Sequence

Given our holistic Interaction Model derived in Step 2, we are unable to represent specific aspects of taskwork, such as task structure, timing, and load. To better represent these components, we can show how the mission can be represented as overtime for more dynamic visualization of teaming. As we've mentioned, the interactions that arise from teaming are often variable and shift depending on the context of a particular moment in time. In this step, we show the actions of the mission over time for two humans moving a table. This allows us to also observe how task structure, task timing, and taskload can impact each agent, as well as observe the peaks and depressions of teaming.

We create a simple dynamic visualization of this team in Figure 2.18.

1. Taskload and Task Timing: We first represent the taskload and task timing by allocating the actions to each agent and organizing the task across time. In this case, both humans complete all the tasks in parallel.

2. Next, we use the information for action to action relationships we found in Step 2 to draw each relationship between all actions. Here we show the monitoring and commanding that can be expected as these humans move the table.

3. Now, we can finally show any task structure interdependencies between specific actions to other action through the line weight. As this taskwork is linear, the next series of actions depend largely on completing the former actions. In addition, as the joint action complexity grows (holding table edges vs walking together, we increase the line weight to show more dependency. Interdependencies between these actions are likely high, with potentially higher demand for monitoring one another than for commands.

We can see that the third action, 'walking together to a new location' seems to be the most impacted action step. Through the interdependency representation, there requires high task complexity due to the monitoring and commanding that can occur, as well as any complexity of the task type, as this action is a joint action that requires coordination and collaboration between the two. Designers may take note to be more aware of any potential problems or factors that may impact the taskwork at this particular moment in time.

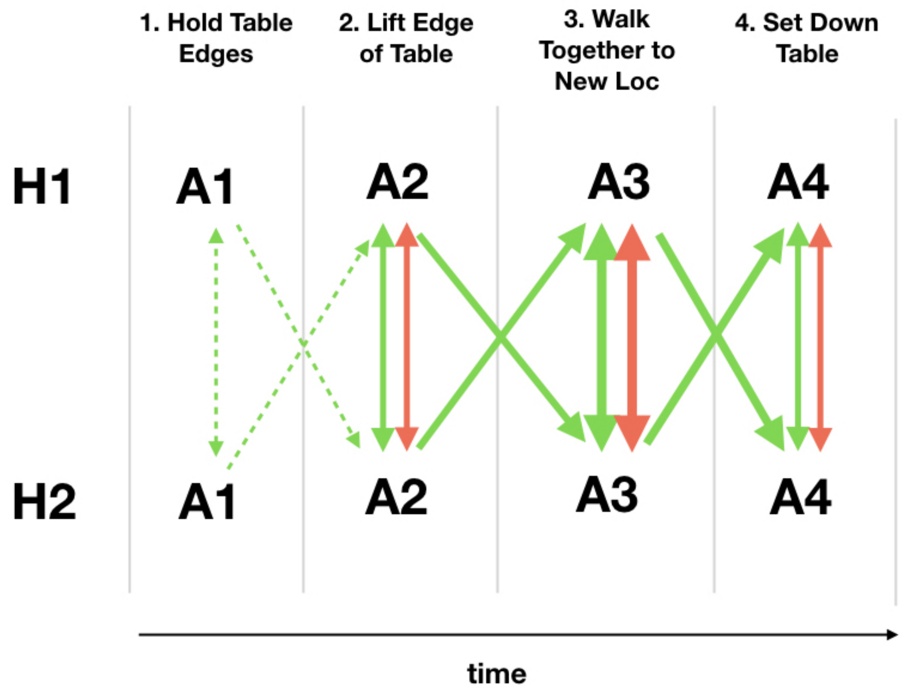


Figure 2.18: Dynamic Action Sequence of Lifting Table Mission

Step 4: Cross Reference with Framework

The last step of the methodology is the mapping of the Interaction Model of Teaming and the Framework of Human-Robot Teaming (Interaction focused) we identified in Section 4. The goal is to determine which interaction components are being focused on and which components have yet to be considered. In step 2, our interaction model shows which areas are impacted by other agents, vs areas that are more self-regulated. We map these GIAPE impacted elements from the Interaction Model to the overall Interaction Framework, see Figure 2.19. These composites should be created on a per agent basis, as every agent has their own internal process. Since both humans are stated to be of similar roles and capabilities, we assume they share a similar mapping and thus show just one mapping.

Here, we can see the coverage of components of teaming interaction for the areas of the cognitive model that are impacted by this mission. Red and green highlighted terminology suggests components that are highly impacted by the design of task complexity, task structure, and task type. Blue highlighted self-motivated terms may be more impacted by the team members themselves with respect to capabilities and skills. While we can see the highlighted portions, combined with the breakdown of teaming in Step 1 and information held in the Interaction model, designers can take these observations into consideration. At this step, designers may find new concepts to focus on for team design, or confirm they have targeted concepts that are critical to the team.

This suggests the highlighted concepts are key areas for team designers to be aware of. Alternatively, it also shows the concepts are may not be focused on. Ultimately, it is up to the team designer to filter through these components and determine what is most important to the team design and performance.

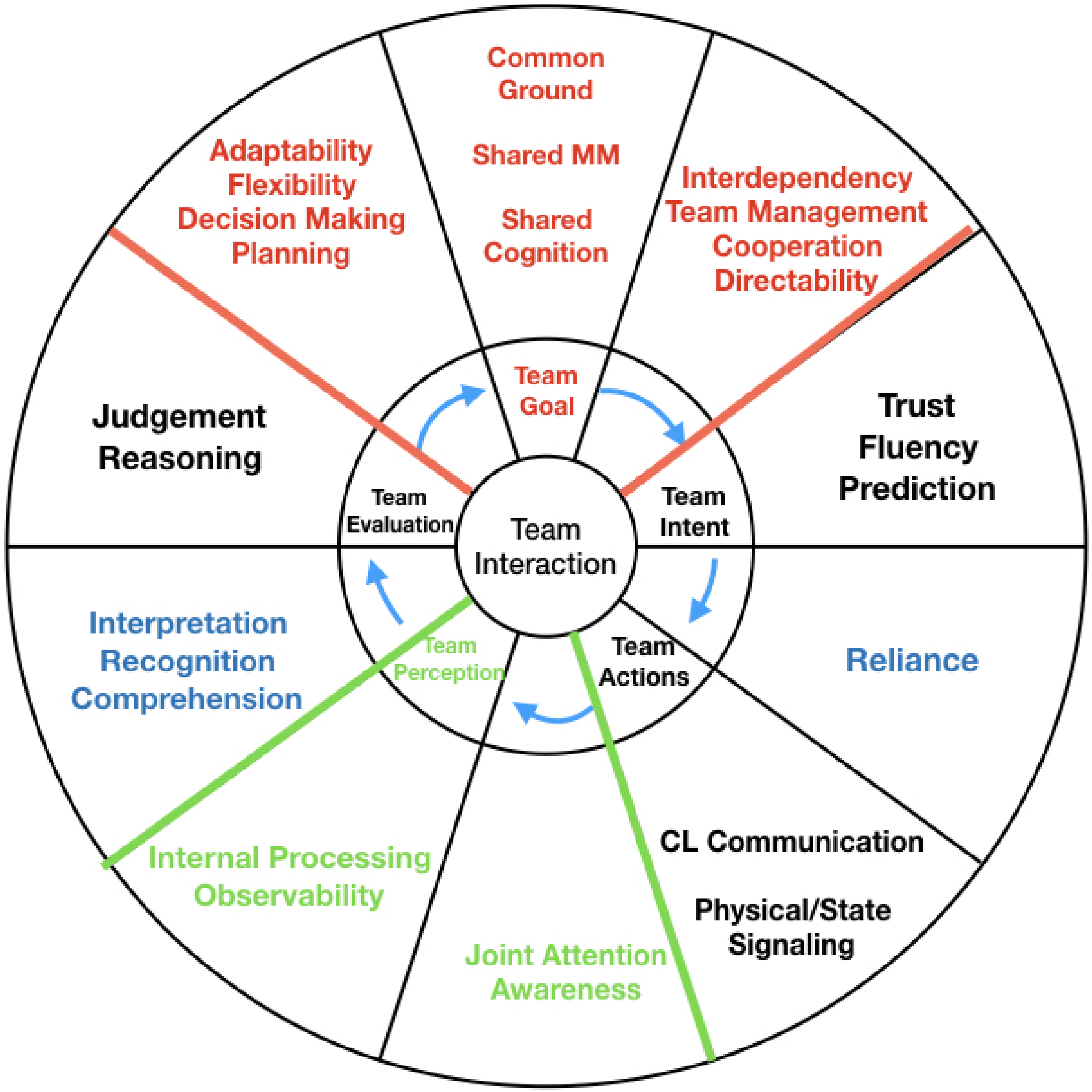


Figure 2.19: Comparison of overlap mapping with the Interaction Framework Concepts.

2.4.2 Verifying the Framework with Previous Case Studies and Scenarios

Now that we've walked through a simple example without methodology, we can show different examples of how to apply it by selecting case studies built of real human-robot teams in the literature. These distinct teams will show the similarities between the key components of consideration found in the results of these case or field studies, and what our framework of human-robot teaming suggests as key concepts to consider for designers.

By basing these case studies of real teams, we can demonstrate usefulness in identifying hypothesis or useful components to assist in making a design decision for the team.

We will explore the following three case studies: a human SWAT case study to imply robotic design, a rover recon mission on a planetary surface, and a future-day in-orbit panel and filter maintenance mission. Table 2.1 describes a high-level overview of each teams similarities and differences.

Table 2.1: Differences and similarities between the three case studies.

Options	Detail	SWAT	Recon	In-Orbit
	Joint			✓
	Close	✓	✓	✓
Task Type	Remote	✓(most)	✓(most)	✓
Team Structure		spread out (fan)	chain of command	network
Info Transfer		constant w/ minimal delays	high delay time	minimal delays
Team Composition		all human	2H:1R	2H:2R
	Monitoring	✓	✓	✓
	Command	✓	✓	
Task Complexity	Self	✓		✓
	High	✓		
	Some			
Task Structure	Little		✓	✓

After applying our Methodology to each of these teams, we will provide an analysis of teaming components our methodology implies compared to the findings of field research of these teams. Step 4, mapping the Interaction model to the Human-Robot Teaming Framework, will be analyzed altogether for all teams later on.

SWAT Case Study Jones et. al. observed field operation exercises of Special Weapons and Tactics (SWAT) teams to determine how to build robotic peers to be tele-operated and used in similar situation [121]. While the team is human-only, this study to suggested how a future-day robotic agent could be designed and which factors of human-robot teaming could be kept in mind.

SWAT teams have a few top-level mission controllers including tactical commanders, hostage negotiator lead, and logistics supervisor, but we will focus on the field tactical commander (TC) as the mission supervisor. These TC work with teams of snipers, scouts, and emergency response agents (ERT), by which we will refer to each team as a single unit for clarity. The TC works remotely from all teams and communicates through voice only; the snipers, scouts, and ERT do no communicate directly with one another.

1. Step 1:

   External Factors : because this is an exercise for the SWAT team, much of the unknowns in a true hostage rescue mission are non-factors to the designers. However for the sake of practice, to the individuals doing the exercises, there may be some unknowns that arise as they progress through the mission.

   Team Members : One tactical commander, a group of individuals who are snipers, scouts, or ERTs. For this example, we will refer to each type of member directly, despite there being a group of individuals that form that grouping. Therefore Snipers refers to all individuals that are part of the sniper sub-team.

   Team Structure : The TC is the top level supervisor who remotely organizes and commands the team of snipers, scouts, and ERTs. These three sub-teams are peers to one another and work in the same close space but do not communicate directly with one another. While the team is all human, the sub-groups are not allowed to deviate autonomously from the TC's commands.

Taskwork : We investigate a hostage situation, in which three teams on-site (scouts, snipers, ERT) work with the remote TC to find and rescue the hostage.

(a) Get and inform mission to team

(b) Determine area layout

(c) Search for hostage

(d) Rescue hostage

(e) Capture target

Task Type: Snipers operate remotely from Scouts, and ERTs, which typically operate remotely or closely with each another(specific cases when ERTs must go on site where Scouts may be). All three groups operate remotely from the TC.

Task Complexity: The nature of these tasks are high risk and require consistent commanding and monitoring from the TC to each respective group.

Task Structure: Despite the TC being the head of this team, all actions are highly interdependent on the results of each group. There is a playbook by which the tasks in the mission should be executed but the team may operate based on the command of the TC operator.

2. Step 2: Creating an Interaction Model of Teaming Given these descriptions of the team's factors, we can first define the action to action relationship via the taskwork components just identified, see Figure 2.20.

   (a) Task Type: the spatial proximity is remote from TC's actions to the other agents (dotted). Snipers are remote from the other agents as well (dotted).

Scouts and ERT can be in the same location but not necessarily working together on the same task all the time (share space, dashed).

(b) Task Complexity: requires TC to directly command (red) and monitor (green) snipers, scouts, and ERT. Meanwhile, there may be some monitoring between the snipers, scouts, and ERT if they interact in the same environment. The green * line between scouts and ERT refers to potential monitoring of each other's actions if brought to the same location by TC.

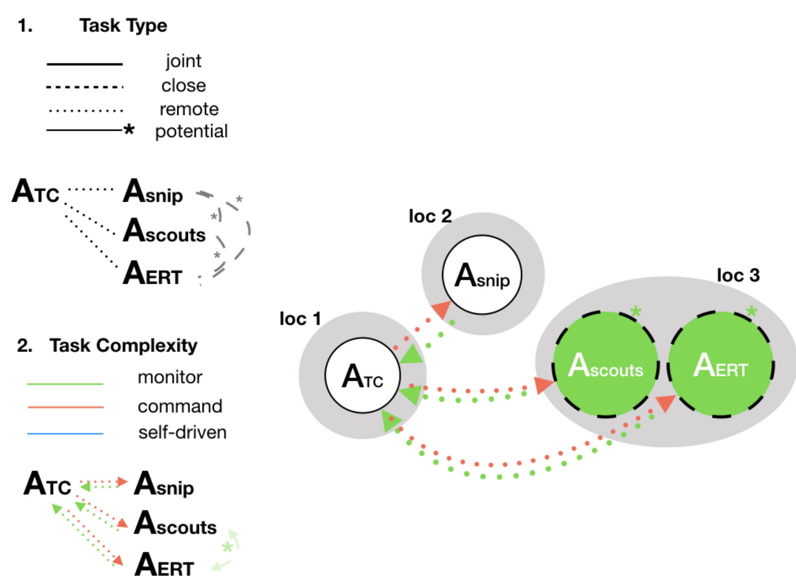


Figure 2.20: Creating the Model for the Military Mission

Once we have this conceptual model of actions between agents, we can construct the entire Interaction Teaming model, see Figure 2.21. Team Structure here implies a hierarchy of TC to all others, as TC is the overall operator. It corresponds to high interdependencies between the GIAPE elements of the TC to the others and vice versa. The mission cannot continue until commands from TC have been met and confirmed by the rest of the team. The other three subunits act as peers but independently from one another. Task Structure implies dependencies between communicating what has been observed to the TC, and how the TC evaluates and formulates a new Goal ‘G’ and relates the next series of commands and sub-goals ‘g’ to the sub-units.

(a) Goals: There is a clear distinction of how the TC determines any shifts to the goal based on their understanding of the environment at any given point in time, and how the sub-goals (goals of the subordinates) are motivated by the TC’s direct mission goal.

(b) Intent: All agents form their own intent as human agents.

(c) Action: Here the actions of the subordinates are funneled into the perceptions of the TC to make an informed evaluation of the circumstances, albeit remotely. In turn, the three subordinates are able to perceive the others’ activities if in close proximity.

(d) Perception: TC receives remote monitoring of all agents’ actions. Snipers, scouts, and EMT can potentially* monitor the actions of agents that happen to be in their area.

(e) Evaluation: As we are considering humans only, we know they can process through to their own evaluations of the circumstances. All evaluations of the scenario are directed back towards TC who requires this information before he/she can formulate the next goal. Here we show the TC’s evaluations able to impact the goals of the other three.

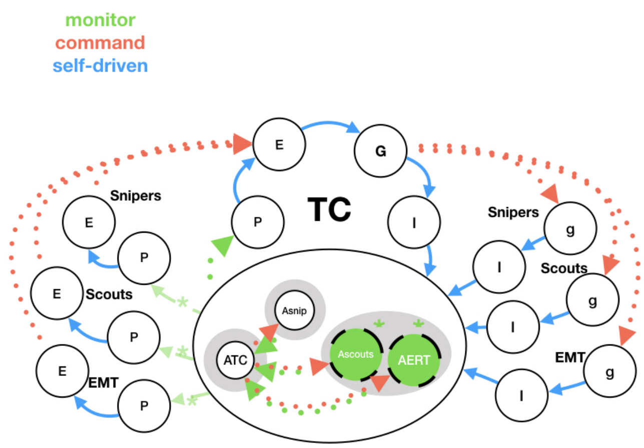


Figure 2.21: The full model of teaming for a military field study.

3. Step 3: Dynamic Action Representation Here, we provide the spread of this mission over time at a high level in Figure 2.22. The first action, 'get and inform mission' propagates down from TC the rest of the team. Afterward for each sub-task, the snipers', scouts', and EMT's actions are perceived by the TC through direct communication. All actions are commanded from TC, and monitoring of each of the groups on the field is only directly accounted for when the scouts and EMT appear at the same location for the 'Rescue' taskwork.

   - Task Timing and Taskload: while generally the overall mission has a linear sequence, the individual actions can be completed in parallel.
   - The action to action dependencies are transferred over to these individual actions. We see the representation of Task Type and Complexity through line type and color.
   - Task Structure: between actions is shown with line width where there are high interdependencies between specific actions.

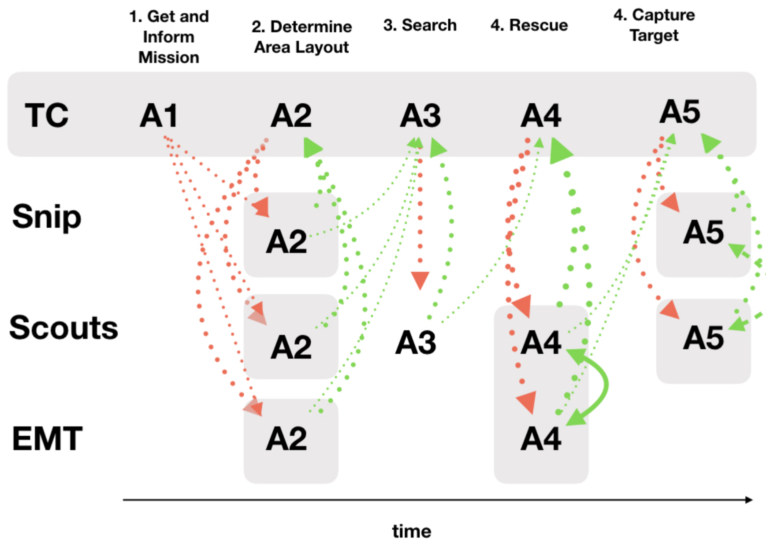


Figure 2.22: The dynamic breakdown of the mission over time.

Rover Recon Scenario The next case study is a rover recon in which a rover will traverse a planetary surface looking for best pathways and science locations [122, 123, 124, 125, 126, 127, 128]. We first assume we are given a satellite map of the planetary location and we reduce the remote ground control into two members for clarity. The goal of the robot is to work with a remote team without direct physical human assistance.

1. Step 1: Defining Teaming Factors

   External Factors : the setting is a planetary surface like Mars or the moon. We will operate under the assumptions of common remote missions like communication delays.

   Team Members : two human agents and one rover robot: FD, the main supervisor for the mission, a human driver (D), and a rover (4 wheel drive, cameras for inspecting surface, lidar for terrain). The Rover can perform autonomous navigation with obstacle avoidance and steer/drive continuously.

Team Structure : The FD agent is the mission supervisor above all other agents. The Driver is a superior to the robot and directly controls it. The Rover has some capabilities (navigation, locomotion, localization, panorama acquisition) but remains largely controlled throughout the majority of tasks. At any point, the ground can do manual control.

Taskwork : The rover recon mission involves four main steps: 1) generating a map of a mars or moon planetary surface, 2) a rover using this map to identify areas of operational risk (steep slopes, rough terrain, steep drop-offs) and answer some science operations questions (whether contacts between units are visible or not, accessibility of units of interest), 3) generating a high fidelity map with information from the rover and 4) a human EV operation to complete scientific experiments and exploration on the mapped out planetary surface, [122]. The example below involves the team performance at stage 2: to traverse and explore the given map and determine obstacles, risks, and answer scientific questions.

(a) Mission setup

(b) Check path to science station

(c) Localization, accessing imagery, location, and orientation.

(d) Navigation (local) to waypoint

(e) Execution of science experiment at waypoint location

Task Type: FD and Driver sit in the same closed environment which is remote from the rover operations.

Task Complexity: While the robot remains somewhat autonomous and able to receive instructions, the overall mission is highly monitored and controlled by the Driver who then reports back to the FD. Any top-level changes and commands come from FD and trickle down to the Driver and then Rover.

2. Step 2: Creating an Interaction Model of Teaming Given these descriptions of the team, we can create an action-action relationship model, see Figure 2.23.

   (a) Task Type: Spatial proximity is remote between the FD and Driver to the Rover (dotted). Here we can see the interaction between FC and driver is direct (solid). There does exist the potential action relationship between FD and the Rover in specific cases such as emergencies.

   (b) Task Complexity: FD directly commands to the driver who then commands the rover remotely. Conversely, the rover's actions are monitored by the driver, who is monitored by the FD. The green * line between rover and FD implies potential monitoring of the rover's actions if direct access is needed by the FD in a given scenario.

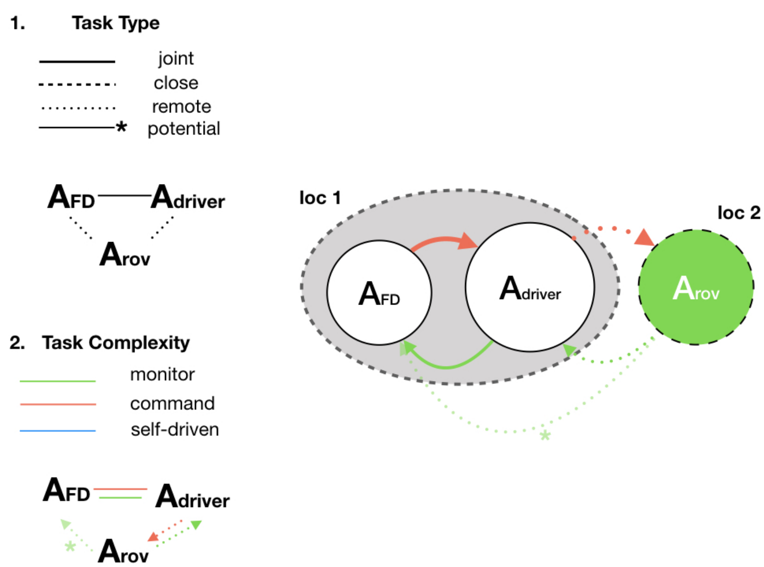


Figure 2.23: Creating the Model for the Rover Recon Mission

To create the giape model of interaction, we formulate the following definition of this figure **??**. Team Structure here is hierarchical and apparent in the connection between the goals 'g'. Task Structure places high interdependencies between FD to the others and vice versa, as the chain of command trickles top down through the GIAPE elements.

(a) Goals: FD's general mission goal gets passed through to the driver. The subgoal 'g' is passed down then from the driver to the rover through direct commands.

(b) Intent: The human driver and FD formulate their intentions based on their goals, but the rover's intent, is remotely commanded and therefore influenced by the driver.

(c) Perception: Here we see the perceptions of the driver's and rover's actions feed directly to the FD supervisor. FD remotely observes the rover and directly observes the driver.

(d) Evaluation: The rover's evaluations of its taskwork goes back to the driver, who then relates this back to the FD at the top. These evaluations of the environment and their current states impacts the top level decisions that FD may make.

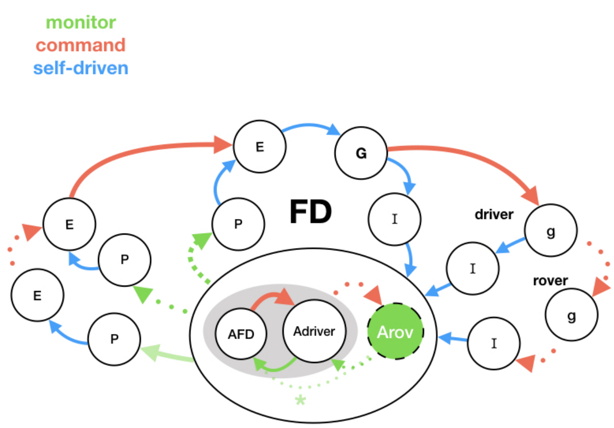


Figure 2.24: The full model of teaming for a rover recon mission.

3. Step 3: Dynamic Action Model of Teaming In this team's dynamic action

breakdown, we see a different pattern than the SWAT team's which was more of a 'fan out' delegation of work. Here the commanding of actions flows down the hierarchy of FD to the driver to the rover, while monitoring of these subsequent actions travels up the chain of command, see Figure 2.25. While FD may subtly monitor the rover's action when needed, it is the driver's main responsibility to monitor and track the rover. During peak points of critical activity like localization and execution of the experiment, the FD may monitor the activity more. These critical tasks also create more interdependencies between agents who's tasks rely on the prior agent's. We add in the action-to-action dependencies (task type, complexity; team structure) from Step 2 to this depiction of the mission through time.

- Task Timing and Taskload: The main task for FD is commands based on observations as the mission progresses. The driver and rover are intimately enmeshed through their joint work, despite the remote nature of this work. Each task completed by the robot eventually propagates back up to the FD through the Driver.
- Task Structure: Independencies here span directly from action-to-action. Here task 5, Execution of experiment, is a task with high dependencies on the localization and navigation prior. It is also interesting to note that all tasks within these five task descriptions are required to progress linearly. This is reflective of the strict and calculated nature of space operations.

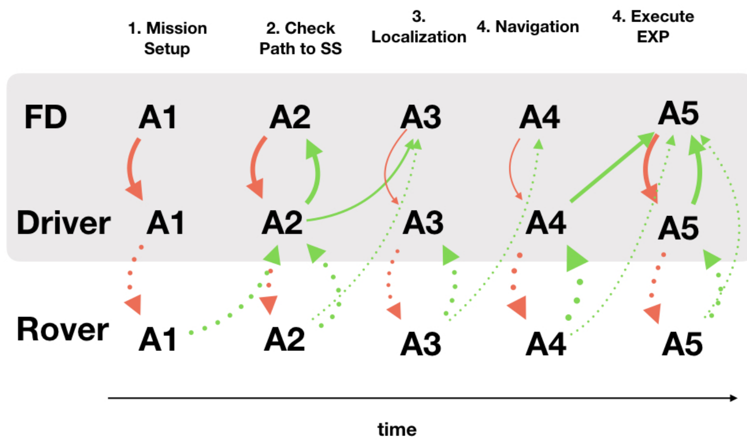


Figure 2.25: The full dynamic breakdown of the mission over time.

Future-day In-Orbit Maintenance Mission In this last case study, we observe a scenario imagining a future day mission for a human-robot team conducting routine maintenance on the exterior of the spacecraft. This team scenario has been explored through modeling and simulation and components of this mission have been undergoing validation through human-in-the-loop experiments [118, 52, 51].

1. Step 1: Defining Teaming Factors

   External Factors : the setting is on the exterior of a spacecraft while in-orbit. We imagine the team working away from mission control autonomously as a unit. In this scenario, the modeling and simulation abstract away external factors that may impact the team by random chance or interference.

   Team Members : two human agents (IV astronaut and EV astronaut), a Free-flying robot, and an RMS robot. The Free-flying robot is capable of assisting the IV astronaut with visuals and inspections of panels.

   Team Structure : The ratio of humans to robots here is 2:2. The IV astronaut remains remote throughout the mission and provides commands and controls for the two robotic agents. EV does not explicitly receive commands to do work from the IV but does have guidance on the mission goal from IV. EV works directly with the RMS and closely with the Free-flying rover through communication. The robotic agents can be controlled directly or through commands by human operators; this assignment depends on their capabilities for the action in question and allocation of the taskwork to an agent.

Taskwork : an extra-vehicular mission in outer space, in which humans need to work together with various robots to maintain the spacecraft. The team inspects ten items (three space panels, four solar panels, and three filters) and replaces them if needed. The tasks to be conducted for the maintenance mission are the inspection of the space vehicle's exterior at three locations. At each location, the condition of the exterior panel needs to check, as well as the wiring and filter underneath the panel. If it turns out the condition of these components has deteriorated, they need to be replaced or repaired. The scenario was set up in such a way that each location needs once repair or replace task (in order, panel replacement, filter replacement and repair of wiring).

(a) Prepare and leave the dock

(b) Gather inspection tools

(c) Inspect Panel

(d) Unscrew Panel

(e) Lift Panel

(f) Dispose Broken Panel

(g) Bring new Panel

(h) Emplace Panel

(i) Screw in Panel

Task Type: The tasks here range from simple traversal and observation to demanding and physical. Tasks conducted by the IV astronaut are remote, while all other agents engage in close or joint taskwork.

Task Complexity: The complexity of these tasks and the operations itself requires close collaboration between agents involved as well as coordination when handing off different components of the tasks to other agents.

2. Step 2: We describe the Action-Action relationships in the following and Figure2.26.

   (a) Task Type: The spatial proximity is remote between the IV and all other agents. EV, RMS, and FFR are working together in the same general location. However, EV and RMS are the only two who have direct monitoring of each other's actions, as they are required to complete joint taskwork. FFR, RMS, and EV all have the opportunity to physically interact with one another via the * notation.

   (b) Task Complexity: Many actions will be commanded from IV to the others. IV will need to also monitor the RMS and FFR robots directly but does not need to monitor the EV who is able to act independently (other than determining the team's goal). EV and RMS work directly together on the taskwork, resulting in mutual monitoring. RMS and FFR pairwise and EV and FFR pairwise do not directly interact but may collaborate loosely on related taskwork.

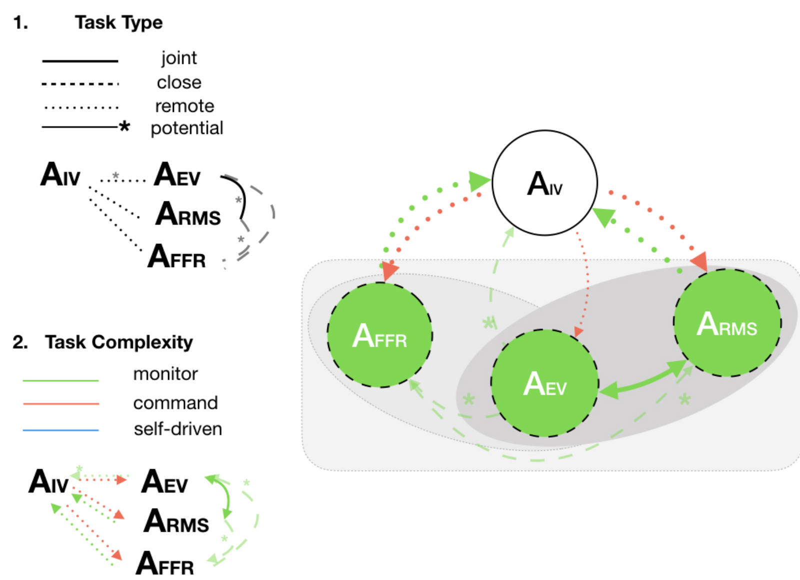


Figure 2.26: Creating the Model for the Rover Recon Mission

Following this, we again generate a model of interaction for this team in Figure 2.27. Team Structure within this team is highly dynamic and complex due to the variety of roles and tasks involved. IV acts as the leader, directing and in control of the overall mission progress. EV is a peer to the robotic agents, working directly with the RMS. The Task Structure for the model has high interdependencies between the IV to the robots, and less so to the EV astronaut. Since EV and RMS work closely, the monitoring interdependency is high between these agents.

(a) Goals: Here we see IV remains as the mission goal 'G' determinant and distributes subgoals to the other three agents. The goals for EV are more flexible as they are expected to be knowledgeable and aware of the mission taskwork. Any evaluations EV makes also feeds back to contribute to their own sub-goal throughout the mission.

(b) Intent: Each agent is capable of formulating intent, as these robots are simulated with high-capabilities. Specifically, this study envisioned FFR and RMS, not as subordinates and therefore able to explicit impact their intent towards a task.

(c) Perception: the perceptions between RMS and EV are direct, whereas the perceptions of other agents are remote or close at best.

(d) Evaluation: While evaluations of the EV and robots' perceptions are directly looped through to IV, EV is also capable of taking its own evaluations into consideration.

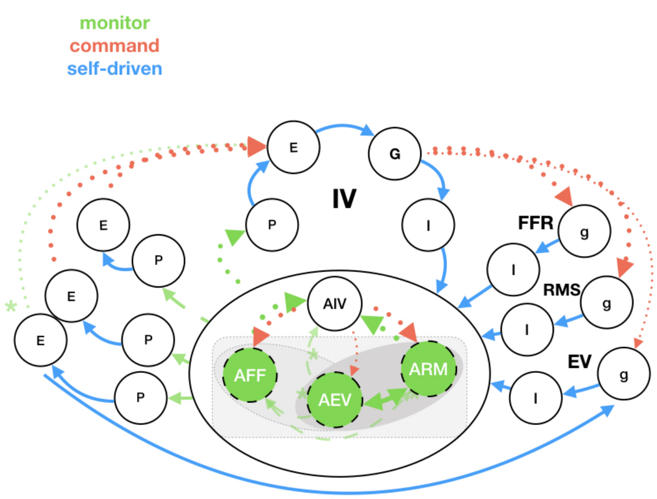


Figure 2.27: The full model of teaming for a rover recon mission.

3. Step 3: Dynamic Action Model Given the most complex action set, we have nine distinct tasks in this dynamic breakdown over time. We observe one loop of this, as it would be repeated during the inspection of each panel. At the first action, IV delegates activities to the rest of the team. Afterward, IV steadily monitors and commands the robot agents, while EV may only occasionally be observed. However, the dynamic teamwork occurs when the extra-vehicular agents perform joint or closely related taskwork at the Panel 1 location.

   - Task Timing and Taskload: While many of these tasks are mostly sequential, there are specific tasks that can be done concurrently, such as inspecting the next panel while a repair is being done. Additionally, the disposal and fetching of panels could be concurrent depending on the capabilities of the agents.

- Task Structure: For dynamic sequences, the task structure of this mission allows for multiple agents to work in smaller teams (EV and RMS, IV and FFR, etc).

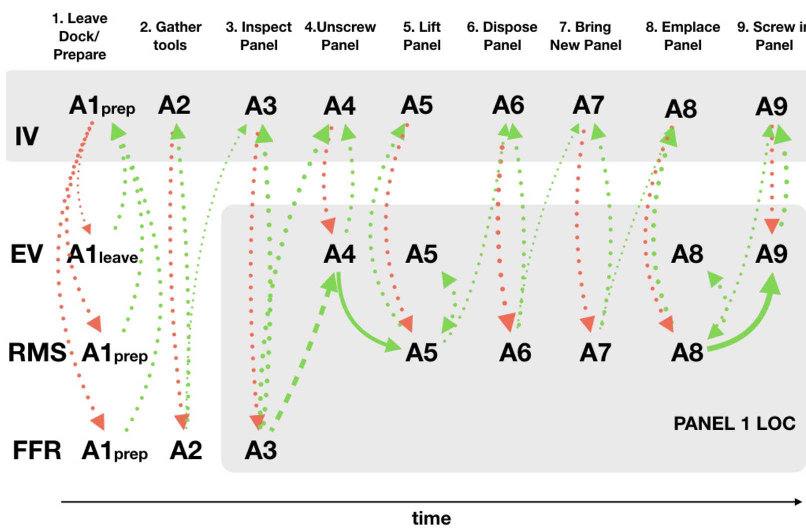


Figure 2.28: A full dynamic breakdown of the mission over time.

2.4.3 A Comparison and Analysis of the Three Teams

After observing these four teams in action, we can compare each one's result of Step 4 of the Methodology: Mapping the Interaction Model with the Human-Robot Teaming Framework. Figure 2.29 depicts all four teams results on the HRT Framework. Since each individual has a unique set of constraints and activity, each team member will have a subtly different mapping. This is especially so for agents that are more distinct. These models should be therefore defined for each perspective. However, for brevity, we will combine each agent's perspective with colors for distinction. We compare and discuss how the distinctions between team inputs from the foundations of teaming impacted which interaction components are highlighted.

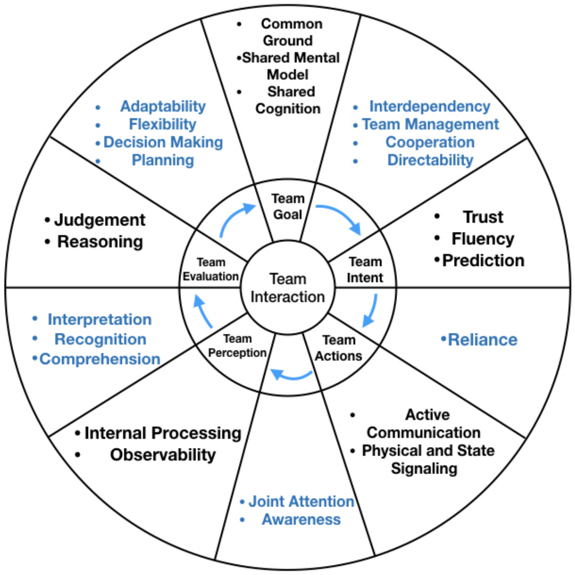
- Team 2 SWAT Figure 2.29b: Given the very strict team structure and clearly defined roles and hierarchy within this team, the Goals and Actions specifically are commanded by TC. While all the individuals involved are human agents, we assume the intention formulation is inherent to each individual. The evaluations of their perceptions are also key components of interaction as these thoughts should be delivered back up to the TC. Perception of the actions of one another and the TC of these three agent teams is another important distinction. Therefore here we know the awareness, observability, and comprehension of the team members is key to delivering accurate evaluations to the TC.

  Jones's paper detailed this SWAT scenario and two significant components of interaction for these types of teams as common ground and awareness [121]. These terms are highlighted in the Interaction Mapping, but there are plenty of more interaction components that could be important to SWAT Teams. As the environment is generally high stress, when formulating and updating the goal for each sub-team, the adaptability, flexibility, decision making, shared mental model, team management, cooperation, and directability are all important. As the team moves towards and completing actions, it may require high trust, monitoring, active communication, and signaling. Within common ground itself, the leadership skills and team management capabilities are important. For awareness, attention an active communication are complementary components to stay aware of.

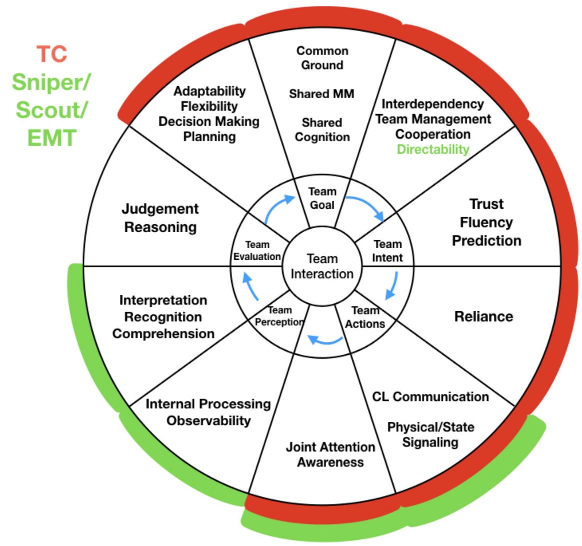
- Team 3 Rover Recon 2.29c: While the taskwork, in this case, may not be as cut and dry, the limitations of the robot's capabilities to operate autonomously impact this framework. Additionally, this requires more commanding and monitoring from the responsible human agents. Here the border for human commands (solid red) stops before intent, while for the robot (dotted red) stops right before actions. In this case, the robot cannot actively formulate intent while the human agent is literally controlling its movements and actions. Within the breadth of rover recon mission descriptions and testing, we found awareness, communication, flexibility, adaptability, and team or rover autonomy to be key components of interest [122, 123, 124, 125, 126, 127, 128]. These components also happen to be highlighted in the Interaction Mapping, and we can address even more components of interaction that may be of interest to rover recon mission and team designers. For the FD, team management, reasoning, and shared mental models are important as a leader in the team. The Driver may need to also be more aware of cooperation, observability, joint awareness, and attention. The Rover and Driver may both need to be aware of observability, active communication and signaling, interpretation, and decision making.

- Team 4 In-Orbit Maintenance 2.29d:

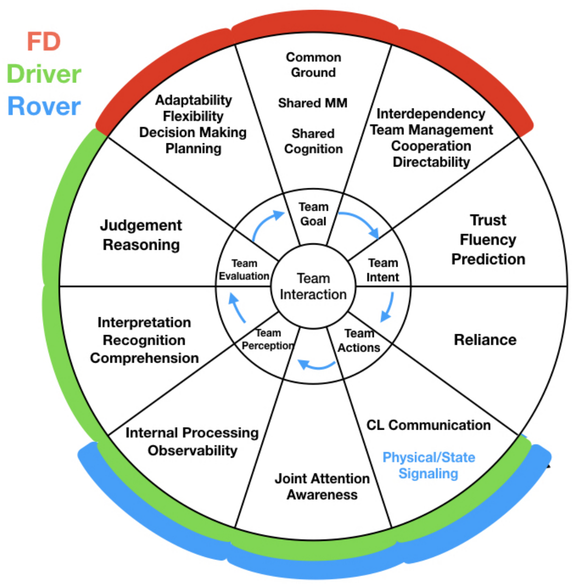
  In this model, the Goal formation, Perception, and Actions stand out. While the team operates under a semi-strict structure, the shared mental model is most important. From the various papers discussing this team scenario, we learn that key contributing factors span not just interaction, but also a team member, structure, and taskwork aspects of the framework. Within interaction specifically, IJtsma and Ma point out the need for understanding interdependencies as well as the ability to handle failure scenarios with adaptability and flexibility [118, 52, 51]. In our Interaction Mapping, we see that these terms and others are highlighted as significant to interaction. The team leader needs to make sure all agents are aligned as there is much trust involved in allowing the robotic agents to have more freedom to act. Additionally, along with the components from the research, other highlighted terms such as observability, predictability, and directability fall directly in line with interdependency needs while shared mental models, common ground, cooperation and more likewise pair with being able to address potential failures. Along the same vein, active communication and joint attention and awareness are important for agents engaging in joint or close activity and communicating physical state.



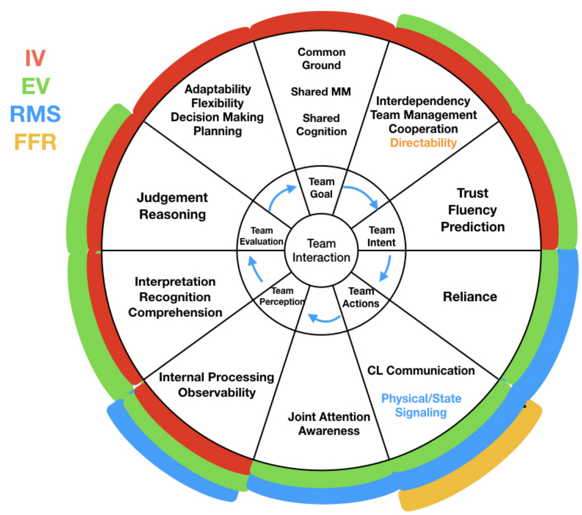
(a) The Interaction composite of terminology on the expanded Scholtz HRI model.



(b) Team 2 Interaction Model Map



(c) Team 3 Interaction Model Map



(d) Team 4 Interaction Model Map

Figure 2.29: A comparison of the differences of Interaction Mapping for the SWAT team, Rover Recon team, and In-orbit Maintenance team.

Now that we can see how the inputs to the human-robot teaming can impact framework across these teams, we can better understand patterns to the dependencies between inputs to interaction. This information can serve to better support designers understandings of the teams they formulate and test, as well as any key interaction components to consider. At a high level, we know that identifying the constraints and assumptions within a team play an important role in setting up the team. But how exactly do the framework components impact these teams?

External Factors are capable of impacting teams with random failures or interruptions. While we did not dive deeply into how a range of external factors can impact a team, the best way to prepare for these components is through planning and awareness. Additionally, designers may want to fully understand the limitations and constraints of their team to understand where any potential weaknesses may lie.

For Team Members, much of the impact that was derived from changes in team members in different team configurations was dependent on each agent's individual capabilities. Agents with fewer abilities and responsibility were more likely to need more support from others. This creates more dependencies that arise when agents of different or similar capabilities work together. Essentially a key factor in team members is understanding how different are the members within the team, and how much support will each agent need.

While all the sub-topics of team structure are impacted interaction, the hierarchy and roles of team members are critical. The number of individuals on the team also adds an extra layer of complexity as the 1:1 or 1: many interactions can occur. These aspects of team structure define just how strict or ad-hoc a team is, and how well they will operate. Less structured and hierarchical teams are more likely to be chaotic and dynamic. Thus designers of such teams need to be aware of this and plan for dynamic interaction to occur as team members interact with one another and their environment.

Taskwork we find to be a monumental component in defining what kinds of interaction will occur. We saw the impact of the taskwork sub-topics trickle down throughout the methodology. Here the kind of taskwork, the complexity, taskload, structure, and timing all change how any agent(s) will work on and complete it. Task structure and complexity even change the relationships and connects that form between different agents. Here it is pertinent to understand the minute adjustments and definition of mission taskwork, and especially the assignment of taskwork to agents. The assignment of work to agents, depending on the prior factors like their capabilities, the organizational structure, and more, will change how an agent can perform a task.

2.5 Future Research for Human-Robot Teaming

While this work provides a first look at the multitude of components that are important to human-robot teaming and interaction, there is still much to explore in this substantially diverse field. We begin to ask questions like what are important factors in interaction for human-robot teams, but also why and how do they impact these teams externally and internally. We hope team interaction designers will be able to use this work to further understand and improve various human-robot teams.

Supporting the design of effective human-robot teaming has always been important to team designers, and future work should continue to reinforce this objective. While some human-robot teams perform day-to-day activities, other teams success or failures provide high impact in extreme scenarios and fields like military, science, and surgical. In particular, as NASA is working towards future-day Mars missions, understanding and supporting the integration of humans and robots in teams will be key to mission safety and success. Specifically, future-day deep space missions will require human-robot systems to be well integrated to successfully accomplish mission goals. This work begins to help solve this problem by providing a series of frameworks and tools to examine teams early-in design.

An important application of this framework is informing the design and implementation of robots for human-robot teams. The Interaction Framework (Fig. 2.9) makes it possible to identify the key components of teaming needed for a particular task. This can inform architectural mechanisms that will be required to support these components and can help software developers by highlighting the critical areas of focus. For example, if the framework for a particular task shows that shared mental models and active communication are important for the task, but prediction is not, then designers can focus on the critical components without needing to implement others that are less useful. The teamwork model (Fig. 2.21) can further aid in this design process, as it represents the main states of each agent and how they interact. This can inform the interaction requirements of the robots, such as the need to communicate remotely, monitor one's teammates, or take verbal instructions from a superior. It also highlights the minimal architectural requirements needed to support teaming. That is, the robots will need to represent their own and others' goals, intentions, actions, perceptions, and evaluations in some formal way. They may also need to share some of these internal states with other agents on the team in ways specified by the model. Finally, the dynamic action sequence (Fig. 2.18) can help specify the joint task and how each agents' actions are related to this task. By representing the task, subtasks, steps, and interdependence relations in this format it provides not only a guideline for designers but also a blueprint for taskwork that can be embedded in the robot software as a way to represent task knowledge.

As we've observed throughout this work, human-robot teaming is a rich area with many facets that have a need for further investigations. Though various research areas have dived deeply into teaming and we have attempted to frame these perspectives into one, there remains a lack of consensus on human-robot teaming concepts and evaluation methodologies. Team designers require better approaches and processes to study and interpret human-robot teaming given numerous scenarios, taskwork, and objectives. Understanding human-robot teaming involves a lifecycle of design, support, and analysis. While empirical studies help confirm and validate real-life scenarios, there is much room for more conceptual and methodological developments and techniques.

While we do explore an abundance of components, we have not touched on social aspects of HRI or more conceptual components of teamwork like cooperation, collaboration, and coordination. These concepts are difficult to define (as there is a lack of consensus) and more difficult to internalize. While these concepts are often mentioned and even assumed to be a result of effective teamwork, they are not necessarily clearly agreed upon. However, throughout the literature, these components have been shown to be useful to teaming and an important factor to achieve within teams [56]. We hope future work can begin to make sense of these more ‘abstract’ notions of teaming.

The terms identified in this work are a subset of and reflect the current status quo of human-robot teaming. As human-robot teams become more advanced, it is likely other human-human teaming terms will become more applicable and relevant. Future work should continue to identify and categorize key concepts for human-robot teaming.

2.6 Closing Thoughts

Human-robot teaming is complex and is rooted in many interdisciplinary fields such as cognitive psychology, human factors, cognitive engineering, human-robot interaction, and robotics. By taking components from human-only teams and human-automation/machine teams, we can distill and condense five key components of human-robot teaming: external factors, team members, team structure, taskwork, and interaction. Human-robot interaction, in particular, is a key component to continue developing research on to better design and evaluate these robotic systems. The basis behind this work is understanding interaction as model of internal processes which map to interaction components.

We hope to have shown the various components of human-robot teaming through a framework that demonstrates the interdependencies among these components. Additionally, by comparing the framework against four different HRI methodologies, we can demonstrate how no one method covers every large factor in human-robot teaming. Finally, we demonstrated how our theoretical methodology can be applied to four different case studies. The comparison of the methodology results in the interaction framework to each case study’s set of interaction components exhibits a need for broadening interaction components for team design. The theoretical methodology is also presented as a way for users to understand the impact these teaming components can have on the resulting team interaction and which interaction components to consider and are important to the team. Specifically, we see the definition of taskwork and task allocation to agents as a factor in the interaction that occurs as a result. We can also pinpoint trends in expected interaction framework outcomes and how specific aspects of external factors, team members, team structure, and taskwork impact interaction in interesting and unexpected manners.

Our goal is to support and exist the development and design of current and future-day teams. We hope team interaction designers will be able to use this work to further understand and improve various human-robot teams. This work will enable designers to dive deeply into foundational components and trends to wholly investigate human-robot teaming and build upon the frameworks established here.

CHAPTER 3

COMPUTATION MODELING AND SIMULATION FRAMEWORK

The basis of this thesis lies with the modeling and simulation tool, Work Models that Compute (WMC). While numerous iterations of this simulation tool have been developed over the past ten years, much of this thesis work has been a process in expanding beyond the foundations of WMC to allow for the analysis of human-robot interaction. This chapter will describe WMC at a high level, the modeling of human-robot teaming components, and the computational simulation behind investigating human-robot team performance [1].

3.1 Work Models That Compute

Work Models that Compute (WMC) is a computational framework that examines the interaction of activity in complex work environments using hybrid-timing simulation. It focuses on modeling the work of multi-agent teams as a collection of resources that define the world and actions that manipulate it. While traditional agent-based modeling and simulation methods represent work as a set of actions that a specific agent performs, WMC assigns actions from a shared work model to agents during runtime. Previous versions of WMC were designed to work with one or two human agents and one or two automated systems in the civil transportation domain [129]. This architecture can thus model a range of different work dynamics to address a variety of research topics [130, 131, 132]. We will go in depth into the conceptual and technical development and additions to WMC in the following sections.

[1]Components of this chapter are taken directly from the papers [118]

3.1.1 Conceptual Model

The following conceptual model will describe the objectives, input parameters, the structure of the model, outputs, assumptions, and simplifications. WMC is written in C++, with additional post-processing analysis written in Python. The current version of WMC used here is a discrete and deterministic simulation, thus all team analysis comes from the input or work allocation to be analyzed. While this work focuses on evaluating different human-robot teams, WMC was originally designed to analyze air traffic management (pilot and auto-flight systems) and other aeronautics-related human-automation teams [52, 118]. The objective of WMC is to quickly simulate and evaluate many different ways to distribute work between agents i.e. work allocation and provide metrics and measures of them.

Content

Content in WMC is the underlying structure of the model and simulation. The three main constructs are actions, agents, and resources. *Agents* perform *actions*, which *get* or *set information resources* or *use physical resources*. The simulation framework applies these three constructs to model teams, i.e. human agents, and robots. Here we will discuss these at a higher level and leave the details of implementation for Section 3.1.2.

The collection of actions, or "work", represent both the taskwork that needs to be performed to complete the scenario successfully and the teamwork required to coordinate the work amongst several agents. The taskwork is modeled as a set of independent actions which can be assigned to any capable agent and are identified through a Work Domain Analysis, a Hierarchical Task Analysis (HTA), or similar others methods. The teamwork actions are triggered automatically when required or may be explicitly modeled similarly to taskwork actions.

Agent models represent the humans, e.g. astronauts and robots that can perform the work. Multiple agent models are available ranging from perfect to imperfect. In this work, WMC employs a human-performance model. The performance model has its own internal dynamics that account for taskload saturation, taskwork delays when agents are busy, and more.

Resources define a range of objects from abstract to physical, therefore representing the state of the environment. The interplay of actions, agents, and resources determines the work dynamics. How the agent performs the task, the availability of resources, and the interdependencies of actions all impact the work and thus the efficiency of the team, Figure 3.1.

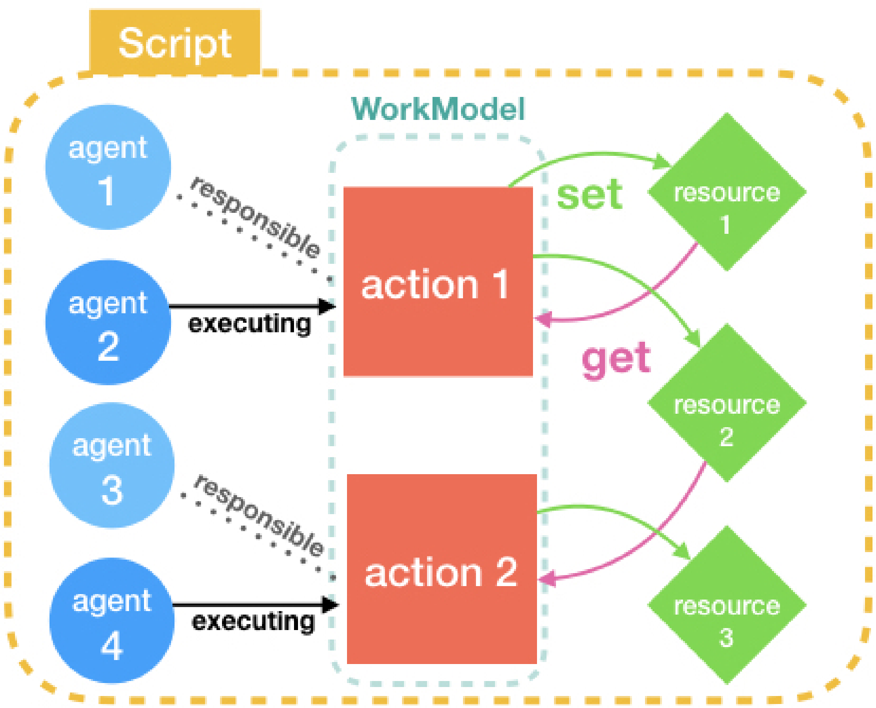


Figure 3.1: The relationship between agents, actions, and resources in WMC.

Input Parameters

Input parameters here are predefined in two main files, one work allocation file, and one roles file. The work allocation file is a list of every action in the mission or scenario, whereby they can have the following attributes:

Authorized Agent the agent who will execute this action.

Responsible Agent the agent who is responsible for the outcome of the action.

Action Attributes determines whether or not there should be monitoring as the authorized agent executes the action by the responsible agent, or if there should confirmation from the authorized agent to the responsible agent.

Action Duration determines how long the action should take.

Collaborative Control Modes determines whether or not the authorized agent needs direct teleoperation, command sequencing, or no control (autonomous) from the responsible agent.

Agent Role a value from 0 to 3, which determines the role the agent will take on in the simulation. The roles range from a controlled robot, robotic peer, human peer, to human supervisor.

Failure Knowledge a boolean value which determines if the robotic failure is known or unknown to the robot.

Failure Instance a numeric value that determines which occurrence of the action should the failure occur. i.e. if 2, then the second instance the action is being executed will result in a failure of to robot to finish the taskwork.

Recovery Executing Agent is the name of an agent who will re-complete the action that the robotic agent failed.

Location is the area in the simulation where the action is to be executed.

Joint Action is a boolean which if non-empty, forces the action to be performed jointly (closely) with another agent which is listed in this space.

Assumptions and Simplifications

Assumptions for this model are generally based on the concept of 'perfect' agents or environments. These assumptions are made to not only simplify the model but also due to lack of data or other information to properly improve or further specify the model. For the purpose of this work the exploration of work allocation is to inform mission planning. The detail of specific robots or astronauts or the configuration of the spacecraft is unavailable or speculative at best.

The biggest assumption of WMC is the concept of an agent model that performs perfectly. In this work, the agent model is not a 'perfect' agent model although it performs actions without failure or delay unless mandated by the constraints of the actions. This basis of using a 'perfect' performance agent primarily based on the logic that even with perfect agents, there will be fundamental problems that can occur within teams. Thus if even perfect agents cannot avoid these issues, then real life teams should expect the same or more in terms of metrics and potential problems within these missions or teams.

Another assumption is that we ignore any external factors that are unaccounted for within the scenario design itself. Along with the perfect agent assumption, WMC also assumes the work is done in a near-perfect environment.

We also assume there are no random agents who can interfere or intervene in the team activity. Only designated and pre-designed agents are populated in the simulation and can affect others.

The general environment of the scenarios we discuss later is simplified due to the lack of external factors or randomness. While these are important to consider, the questions we try to answer do not necessarily require this level of detail.

3.1.2 Software Design and Implementation

The design of WMC is derived from the actions, agents, and resources. This section will describe the modeling efforts for these main components. Verification and validation of these can be found in Chapter 6.

Actions: Work Models

Actions, or work models, in WMC is purposefully modeled independently from any agent such that different work allocations can be easily evaluated. This differs from typical modeling of agents which are object-orientated, where agents have defined methods or taskwork. This particular action modeling allows for the inherent theory behind WMC: to easily allocate various work to different agents.

Work models act as the umbrella or conceptual parent of functions (logical groupings of actions) and subsequent actions (specific components of work). Work models effectively coordinate the actions that belong to it and link agents to actions (allocation of work) and actions to resources (to use, get, or set). Functions define a conceptual 'class' to which a series of actions belong to.

Actions are standalone descriptors of the work that is being performed. Each action has an attribute that defines the interaction with the environment through three types of relationships with resources: (1) an action gets an information resource, thereby retrieving information from the environment; (2) an action sets an information resource, corresponding to changing information in the environment; and (3) an action uses a physical resource, similar to using a tool in real-life.

Actions are modeled with two main methods: a next update time and a resource update method. The next update time specifies when the action needs to updated next. The resource update method manipulates the environment through getting and setting information resources, similar to obtaining information from the environment and making changes to the environment, and/or using physical resources, representing the use of a tool or other physical elements.

Actions have at minimum two agents associated with them: an executing agent and a responsible agent. The former actually "executes" the action, while the latter takes responsibility for the outcome. For example, the task of moving large materials on the dock may be executed by the RMS robot, but the IV astronaut is responsible for the outcome of the task. In the case of joint activity, there is an additional third agent that is capable of assisting with the work. The duration of actions is a value contingent on the agent performing it.

One important result of this executing vs responsible agent is the concept of an authority-responsibility mismatch. This refers to when two different agents are assigned to execute and be responsible for particular actions. WMC employs two automatic responses to this mismatch based on the inputs to action attribute and command sequencing.

The desired control mode, as well as the desired type of verification in case of authority-responsibility mismatches, can be specified on a per action basis [36]. For example, certain actions that are deemed critical may require confirmation by a human before other actions are performed. WMC can engender confirmation actions for these critical cases, whereas monitoring actions can be engendered for less critical actions. Similarly, if the robot is capable of independently performing certain actions (after having received commands), but might need more direct human involvement for other precise or critical actions, WMC can engender command actions for the first actions and control actions for the latter cases. Finally, communication delays can preclude the use of direct tele-operation (control actions) and real-time monitoring, and instead require the use of command sequencing and confirmation actions.

Action Attribute Designers can define either direct (tele-operation or manual control) and command sequencing (supervisory control) to occur when a mismatch occurs.

In the case of direct teleoperation, a control action is engendered that is executed by the human agent. A control action needs to occur simultaneously with the action-to-be-controlled. In case of command sequencing, a command action is engendered that is performed by the human agent and needs to occur prior to the action-to-be-commanded.

Command Sequencing Within command sequencing, a mismatch will trigger a teamwork mode that corresponds to either monitoring or confirmation actions. Upon identifying an authority-responsibility mismatch, the simulation engine automatically engenders monitoring and/or confirmation actions which are then allocated to the responsible agent. Monitoring is a parallel action in which the responsible agent observes the authorized agent during the execution of the main action. Confirmation is a subsequent action in which the responsible agent confirms the successful execution of the main action after it has been executed. The authorized agent needs to wait for confirmation before it can continue with its next action. These monitoring and confirmation actions emerge as the work progresses, and themselves can impose significant taskload on agents.

Agents

Agent models act as a core section of the simulation performance. Agent models do not contain descriptors of specific work activities but rather contain logic to process the execution or delaying of actions. The performance model of the agents essentially calculate or make simple decisions regarding resource updates and measurements. WMC can use any type of agent model that is deemed appropriate for the analysis as long as it meets the computational interface standard of accepting calls from the simulation framework to execute the actions it is passed during run-time. Examples of agent models currently used in WMC are a perfect agent that can execute all tasks instantly and perfectly, and a more-extensive performance model that also adds elements of task management, delaying and interrupting actions when its assigned taskload reaches limits.

At a high level, the agent model takes in actions it is assigned and executes them. Under the hood, the performance model completes a series of checks before actually executing the action it is called to do. These checks, if triggered, require the engendering of other actions or in the worse case, completely delay or reschedule the actions entirely. This work uses a performance model that tracks its immediate taskload and delays or interrupts actions if it is too busy. The performance model performs optimally in the sense of action timing execution with no delays unless otherwise specified. It controls all action execution and engendering of responses specified in the input file. The model also defines the agent's performance directly as it can directly impact the timing of action executions through scheduling additional teamwork actions (more load on agents) or delaying subsequent actions based on physical resource constraints.

Resource(s) Availability Here, the performance model checks to ensure any resources that need to be gotten, set, or used by the action are not being currently used by another action. If any resources are occupied, the action will be delayed.

Resource(s) Location The location of resources that will be impacted by the action is also important. Any resources that are not at the right location will cause the performance model to trigger a 'fetch' action to gather it.

Agent Availability any agent associated with the physical execution or authorization of this action will require their attention, and therefore if the requested agents are unavailable, the model will delay the action until all parties are available.

Agent Location Similarly for agents who need to complete the action physically, the performance model will trigger a 'traversal' action to move the agent to the correct location in the environment.

Joint Agent Availability/Location For any actions that require a joint agent to perform the action, the performance model will also check their availability and location.

Authority/Responsibility Mismatch As mentioned prior, the mismatch of responsibility and authority can create a host of other actions to be called. This logic is handled here in the performance model.

Failure of Actions When specific actions are scheduled for failure, the performance model will execute a set of requirements prior to initiating the failure. This series of actions will be explained in greater detail later on.

Resources

Resources, modeled through variables, are one of the most simple objects in WMC, but provide a large impact on the performance of the system. Resources are defined in the work model or the script (the denotation of each case study). The set of resources specify the environment within which the agents act and are split into information or physical resources. Information resources are elemental pieces of information, ie. the location of a tool, the amount of oxygen left. Physical resources represent physical items in the environment, such as tools, portable life support systems (PLSS) and spacesuits.

Physical resource models are implemented as structures containing lower-level information resources describing attributes (i.e. physical resource: panel, can contain several information resources about that panel like location, condition, is screwed on, is disposed of, etc.). Physical resources are modeled by computational structures defining attributes important to their use, including containing information on the resource's location, current ownership by an agent, and availability.

The internal structure of a resource in WMC is one top-level resource that essentially keeps track of sub-resources that define its details further. For example, this nested structure is typically seen as a physical resource 'wiring tool' that contains many sub-information resources like 'location', 'availability', or 'movability'. As such, this structure allows agents to use the physical resource 'wiring tool' for a repair, but also check its 'availability' or 'location' if needed. The number of information resources that can be attached to a top-level physical resource is unlimited, and many pieces of information regarding the resource can be easily called. While this nested structure is useful, it's not required all resources be constructed in this way, as standalone information resources can also exist.

3.2 Modeling Human-Robot Teams, Teamwork, and Interaction

To investigate work allocation in human-robot teams, we must model and assess both taskwork and teamwork. In HRI, interaction serves as a core concept to center studies for human-robot teams. In this section, we define interaction as teamwork, the result of splitting the taskwork amongst team members. This section defines the aspects of teamwork we believe are necessary for assessing teaming across different work allocations. We categorize team performance into the three levels we believe are important to consider: cross work allocation, per work allocation, and per agent.

3.2.1 Conceptual Expansion of Taskwork and Team Interaction

Interaction within teams can be categorized a number of different ways. When deriving additional metrics to attempt to capture most of the aspects of teaming described in the literature and summarized in Section 2, we found it useful to begin with those behaviors which were externally observable, e.g. physical proximity, task interdependencies, interruptions, and failures. However, we soon found it necessary to also model constructs internal to each individual agent's own cognition to derive objective metrics that would capture the more abstract aspects of teaming such as coordination and collaboration.

External Modeling here focuses on things that are externally observable, such as system failures and taskwork. Modeling specific definitions of task type, task interdependencies, and teamwork taskwork help simulate a more realistic external impact on teams. This allows us to explore a wide variety of situations by adding failures, interruptions, and task-handoffs into the nominal mission timeline to test the robustness of various work allocations.

- *Action Locations* Actions in WMC are location specific, such that there are low or static times allocated for agents to move from one place to another. Traversal actions allow agents to move from any location to a new location with a duration proportional to the total distance to be traversed. There are three kinds of traverses that can take place:

(1) Normal scheduled traversal to the next location after the previous action is completed.

(2) Traversal to reach a place where an action is scheduled and execute that action.

(3) Traverse after fetching a resource to deliver it at the destination.

- *Spatial Proximity* Team members work together with respect to spatial proximity in three main ways: jointly, closely, and remotely, Figure 3.5.
  - *Joint Work* is tightly coupled work that requires close-knit interaction in the same physical space. Agents are working closely on the same taskwork and typically require high communication, coordination, and collaboration to succeed, i.e. lifting a table, drawing on a paper at the same time.
  - *Close Work* is loosely coupled work that is in the same physical space, but does not require agents to work together as closely. In this case, agents are close in proximity to one another but do not need to directly work on the same task. Taskwork examples include orchestras playing music together or two people building separate furniture in the same room. This is a type of work can include task-handoffs, where work is incomplete work is passed on to other team members.
  - *Remote Work* can be for concurrent or distinct taskwork but is defined by the physical separation of agents. For example, mission control to astronauts on the ISS can be doing tightly coupled taskwork but not in the same space. Another example is a team scavenger hunt where team members can communicate through walkie-talkies but have their own lists of items to find.
- *Fetching Actions* bring resources that are required by any action to its location.

When a resource is not present at the required location, a fetch action is automatically scheduled and assigned to the closest, available agent. The agent traverses to the location of the resource, then fetches the resource, and finally traverses back to deliver it.

- *Teamwork Actions* While taskwork can be defined as specific actions to complete, the division of taskwork across agents creates teamwork actions, Figure 3.2. These actions depend on the relationship of the agents working together and how much assistance a robotic agent may require. The assignment of teamwork agents to agents requires understanding which agents are executing agents (those who complete the physical action) and which are responsible agents (those who are responsible for the outcome of the action). The distinction between regular physical taskwork and teamwork provides insight into any added cognitive stress that would otherwise be unobserved. However cognitive load can factor into mission success and is an important component to take into consideration. Four main teamwork actions are required based on the relationships of the team members:
  - *Monitoring* actions are executed when responsible agents must oversee the completion of work.
  - *Control* actions are when responsible agents need to physically control robotic agents to complete taskwork together.
  - *Confirmation* actions are activated after executing agents have finished their work. This acts as a quality control check that the work was completed correctly.
  - *Command* actions are for informing executing agents the work to be completed.

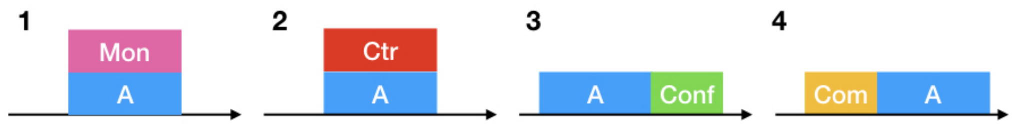


Figure 3.2: From left to right, teamwork actions for Monitoring, Control, Confirmation, and Command actions.

- *Interdependencies* represent the dependencies that taskwork and teamwork have on past, current, or future work. For example, agents may wait for search teams to finish before deciding which area to infiltrate or repair teams may wait for inspection teams to determine which objects need maintenance. This encompasses taskwork that may never need to be completed depending on the prior actions.

- *Failures and Interruptions* Agent failure to complete or do taskwork should be planned and prepared for. However, even the most precise teams cannot predict all types of failures in their systems, particularly ones that are externally driven i.e. robot cannot maneuver around a large boulder. Failures also create interruptions and delays to the workflow, which can impact agents that have interdependent actions. Modeling failures and the resulting agent responses and recovery are interesting interactions that when observed, can yield important caveats or edge cases for behavior and task assignment.

  Robot agents are assumed to be imperfect and may fail to complete specific actions throughout the simulation. As robots lack the ability to communicate through the same subtle cues and gestures that humans do, robots need to communicate their failures effectively.

  Three robot phases of communicating failure are modeled: getting the responder's attention, alerting the responder that help is needed, and creating a request for help. These concepts are modeled as failure script events, interruption actions, re-scheduling failed actions.

Failure script events fail a specific action at a single point in time to represent the robotic failure occurring. Failure events trigger an interruption action and communicate the failure to the action's responsible agent. Robots knowledgeable about failures report the failure to the responsible agent to represent robots getting the responsible human's attention. Failures by robots who do not realize the failure are caught later on.

Interruption Actions impact the responsible agent for the failed action and the agent that will be assigned to redo the failed task. The responsible agent reschedules the failed action to be executed by another agent (human), in figures 2, 3. For example, the Intra-Vehicle astronaut (IV) is responsible and interrupts the Extra-Vehicle astronaut to do the "get new panel" action, which the RMS robot agents failed to do first.

Confirmation Actions are executed by the responsible agent and check the information resource "failure" for the failed action. If completed correctly the mission can continue as normal, Figure 2. If completed incorrectly, there was a failure and this action will schedule an interruption action, Figure 3.

We model two different potential methods of identifying and resolving the failure within the teams, reflecting when failures are either known or unknown to the robot:

If the robot knows of its failure, it interrupts the responsible agent with an interruption action. The responsible agent for this failed action will reschedule and reassign Actions1 as Action1x to be executed by a different agent. This new agent, EV, will complete it. When the responsible agent verifies the relevant agent can continue, the simulation continues, see Figure 2.

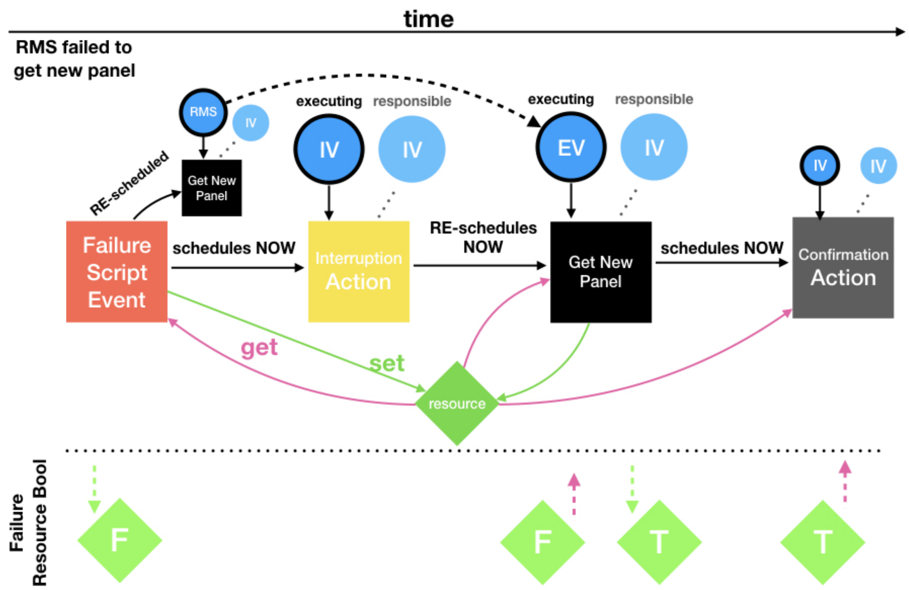


Figure 3.3: The flow of Failure where robot knows of its failure and requests help.

If the robot does not know of its failure and completes the action incorrectly. The failure is discovered later on by the responsible agent via a confirmation action. The responsible human agent must resolve the error by interrupting a new executing agent to re-do the rescheduled action, see Figure 3. Once the rescheduled action is completed, the responsible agent checks that the action has now been completed correctly before continuing.

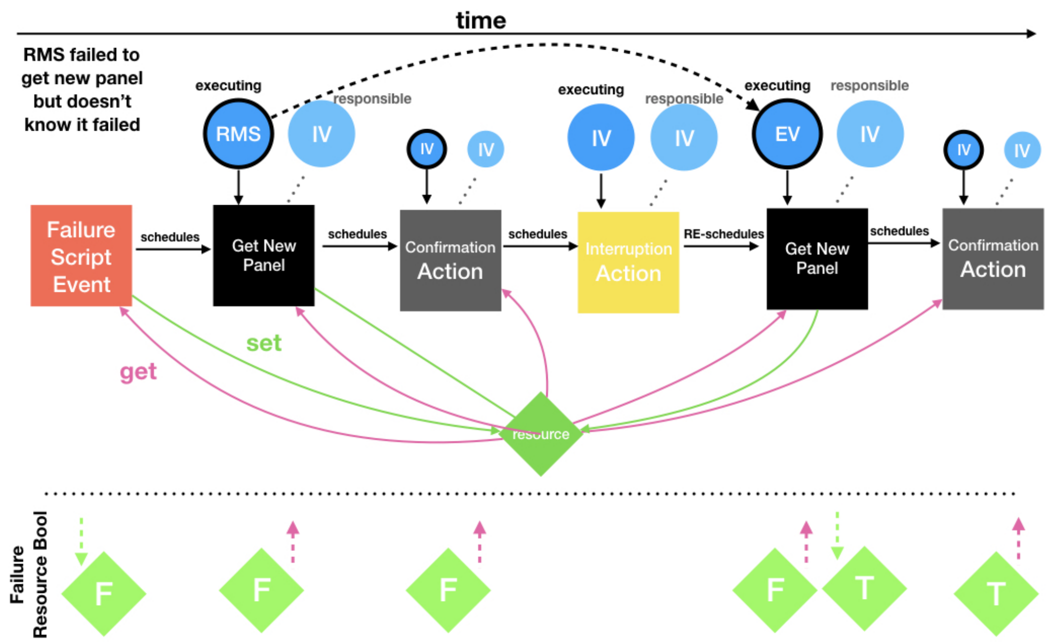


Figure 3.4: The flow of actions where robot fails but does not know of it and does not actively seek out help.

- *Communication* Communication is important to effective teaming, and in our model, we represent this as information transfer. This information transfer is recorded when agents share resources such as information, tools, or GIAPE resources. Specifically, communication of information resources occurs when two different agents get or set an information resource value.

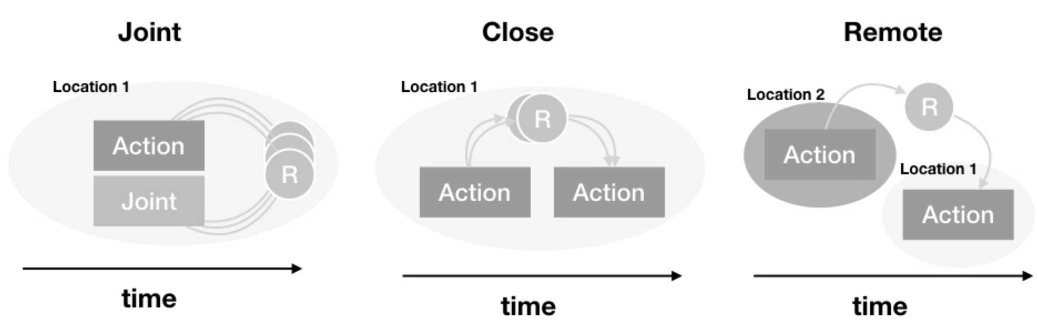


Figure 3.5: Examples of joint, closely-collaborated, and remote work that are defined by proximity and shared resources.

Internal Underlying the externally observable interaction are the internal processes within an agent. Modeling internal processes that impact interaction helps us investigate how these components can affect mission outcome as well as provide a systematic mechanism to model and measure more abstract metrics (coordination, collaboration, etc).

- *Interaction Model* Our approach to capture components of interaction is through a well-known model of interaction which defines the internal cognition of a team member as a cyclic process. Scholtz's HRI model defines this process through goals, intention, action, perception, and evaluation (GIAPE) [84].
  - *Goals* represent what the team and individuals want to accomplish.
  - *Intent* is how individuals intent to satisfy the goal.
  - *Action* is the execution of an agent's intention.
  - *Perception* is an agent's assessment of actions and ongoing surroundings.
  - *Evaluation* is the comparison of a agent's system state to their intention.

  This model expands one view of interaction that focuses directly on the impact of action to action between agents by encompassing internal processing. Shared goals are represented at a higher level than individual actions, while intentions, perception, and evaluation can be updated every action. Actions are already represented in our computational simulation, so, therefore, goals, intention, perception, and evaluation are represented as resources that agents are linked to depending on their respective roles, capabilities, and responsibilities within the team 3.6. We capture the getting and setting of these resources throughout the mission to measure the interaction between agents with the communication of these cognitive process with varying roles.

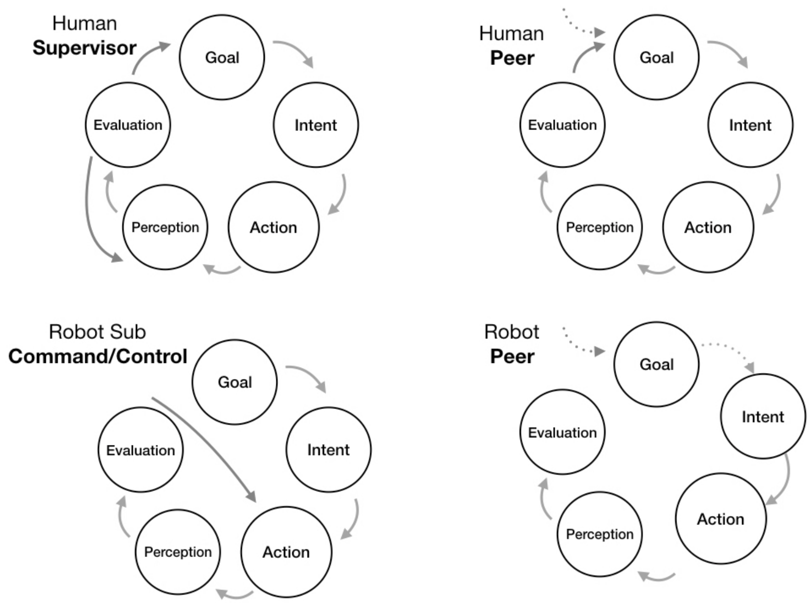


Figure 3.6: Agent Roles for humans and robots and how they factor into capability to impact goals, intent, action, perception, and evaluation.

- *Agent Roles* The concept of agent roles, essentially the function or duty assumed, is tied closely with this cognitive interaction model. The specific role of an agent will in turn impact how they interact with these GIAPE components. For example, a bystander or human-controlled robot's role in a team does not allow them to change the overall goal of the team. This type of authority lies with a team supervisor or manager. Interestingly, while some definitions of roles in teams can be explicit, other times the role that a team member shift over time depending on the agents they may be working with, the specific task, and context of the taskwork. Realistically, agents can take on other roles in different situations and with other teammates. For now we define the agents' roles that each member takes on throughout the entire mission. This is in part due to the specificity and nature of a highly controlled environment (space maintenance). In doing so, we define specifically how agents will work together as they find themselves interacting with their peers, superiors, or directs. We defined four roles that agents in this scenario take on and how these roles define how they impact the cognitive components of GIAPE in a team, see Figure 3.6.

- *Supervisor (human)* roles define agents that are responsible for the general success and workflow of the team. These agents set the goal for the team and confirm the intentions, perceptions, and evaluations of other team members throughout the mission.
- *Peer (human)* roles for humans specifically are generally responsible for other robot peers or subordinates. However, they still check in with the mission supervisor as they progress with the mission and must communicate their intentions, perceptions, and evaluations.
- *Peer (robot* roles define robotic agents that can work with human peers but may require additional assistance with their workload through command. These agents rely on human companions to check and monitor their workload.
- *Subordinate (robot)* roles define robotic agents that have low autonomy and rely almost entirely on human peers or supervisors for assignment of goals and action execution through control. These agents are not capable of higher level cognitive processes such as intentions.

3.3 Closing Thoughts

WMC is a unique modeling and simulation tool. Its components and infrastructure are always changing and updating to model more realistic and detailed human-robot teams. As others continue to advance WMC, the simulations become more and more reflective of human-robot team performance. The ability of WMC to simulate not just human-robot teams, but any team configuration (human-only, human-automation, etc) also provides much design flexibility. Though the idiosyncratic nature of each team configuration, mission, and other design details that are often difficult to capture outside of HITL studies, the construction of WMC is able to capture the important aspects of taskload and team interaction [2].

[2] While computational simulation is useful for providing quick turnaround on potential team performance, it must be validated to provide trustworthy results. The verification and validation of WMC components are specified further in Chapter 6.

CHAPTER 4

EXPANDING METRICS FOR HUMAN-ROBOT TEAMING

Methods to model, measure and predict effective human-robot or human-automation interaction have made significant progress towards designing effective and efficient teams. However, studies of Human-Robot Interaction (HRI) have typically focused on competing metrics of performance and efficiency. Less well understood is how the division of the work between human and robotic team members impacts, and is impacted by, the inherent teamwork requirements that result from this division. Further, different interaction mechanisms between humans and robots determine how this teamwork can occur, further impacting performance efficiency.

Historically, many different approaches have been applied to the challenge of integrating machines into human teams. In the HRI domain, a common approach has focused on predicting specific metrics such as time to complete a joint task or the minimization of idle time [133]. In the human-automation interaction (HAI) domain, team design includes the task of allocating work among agents. Within HAI, historic approaches focused on "function allocation" discussed the high-level division of work according to broad concepts such as roles or generic functions, or according to levels of automation such as those defined by Sheridan and Verplank [53]. More recent approaches have noted the further need for "work allocation", a more detailed evaluation of the specific task work actions that the team must perform [36], and the teamwork required to coordinate the task work allocated to different team members. These metrics span both taskwork and teamwork and highlights the need to consider both [36, 50, 51, 52].

Design of human-robot teams further needs to examine the mechanisms by which humans and robots interact [1, 2]. The following section will address the additional challenges of assessing and measuring human-robot teams by expanding on the allocation metrics developed earlier for HAI. HRI interaction is modeled as including both internal and external processes. External processes include joint responses to failures and other interruptions of taskwork, and also different types of taskwork interaction (close and remote). Models of internal processes represent the cognitive process inherent in interaction (shared goals, intentions, actions, perception, evaluation) and communication within human-robot teams.

Analysis of the resulting metrics identifies relationships between allocations of taskwork, mechanisms for HRI, and interaction within human-robot teams. This expands the set of metrics by which human-robot teams can be evaluated.

4.1 The Context and Limitations of Metrics

While each team differs and may benefit from different measurements, the most impactful results come from finding the best set of metrics, or sometimes even the "metric that matters" [134].

However, even after finding a metric that seems to capture key aspects of the system, researchers may find themselves asking about the meta-relevance of the metric itself. How well does this metric capture the system? How reliable is this measurement? In what conditions is this metric most useful?

To answer those questions, researchers must define the context or scope of metrics, which is key to understanding the caveats and limitations of what is being measured. Metrics are most useful when there is a threshold to compare to. For example, if your metric for total mission time is x minutes, one cannot assume it is good or bad without knowing the maximum time allowed for the mission, or the fastest time possible, etc. Limitations are also consequential when discussing metrics, to understand when a metric may or may not be applicable. If a measure is applicable under specific circumstances, then these conditionals are important to note.

To find the most impactful metrics, we can observe measures at different levels of abstraction. At the highest level of operations are overall mission level metrics, followed by team metrics (and any sub-teams), an agent level metrics (split into human vs. robot). In doing so there can be measures that observe the results of the work or task completed, the success of the team and its members. However, even within these levels, there are be many specific metrics that target particular pain points.

More specific metrics can look at components of team performance that are particularly interesting. For example, team success may be defined by maintaining a steady average of work across the mission or by achieving optimal levels of some metrics while letting other metrics fall to the wayside. Along the same vein, designers may choose to optimize for highs, lows, peaks or valleys in the metric results. In doing so, researchers may find they must make many tradeoffs in determining a set of metrics that matter.

Metrics, therefore, do not necessarily define team success, but rather complement the discussion of what defines a human-robot team's success. While team designers are putting together teams, missions, and other components, they should also be thinking about what metrics will be important as well as their context and limitations.

4.2 A Variety of Perspectives for Investigating Human-Robot Teams

There are many approaches to defining categories of metrics for investigating human-robot teams as prior research has tended to focus on specific aspects of human-robot teams. While these perspectives offer some insight into defining team performance, there is a lack of consensus on what defines an effective human-robot team [50]. One high-level perspective contrasts mission performance with team performance:

Metrics of System and Mission Performance typically refer to the quantitative evaluation of performance and efficiency; this has been also referred to as "common metrics" [100]. This category can also consider subjective ratings. Examples of these metrics include mission performance and execution such as mission duration, mission success, task metrics (duration of actions, task execution), and navigation (obstacle encounters) [100, 135, 136].

Team and Interaction Performance metrics also provide a holistic view of a team but focus on more qualitative metrics, of cognitive aspects of the teams such as teamwork. Different perspectives highlight different key metrics. One perspective distinguishes between *physical interaction* and *cognitive* metrics. *Physical (interaction)* metrics describe physical contact and interaction between team members, such as physical proximity or duration of joint activity. Researchers focused on physical interaction may focus on smaller details such as response time, agent availability, task handoff, [133]. *Cognitive* metrics include information exchange/communication, decision and action selection, and inherent lag or delays. Authority and responsibility relationships between teammates here play a key role in decision action selection and communication specifications [133]. Information assessment and exchange here can also be relative to neglect tolerance, situational awareness, and regulation of control [133]. Cognitive measures and internal processes within agents can also be determined through the understanding of any shared mental models, trust between team members, behavior acceptance, observability of teammates, and coordination. Specifically, some have observed mental model convergence, mental model similarity, and fluency [137].

Another perspective applies to Freedy's Collaborative Mixed-Initiative System. This system distinguishes between metrics of performance (often quantitative and results focused) and metrics of effectiveness (subjective and qualitative) [70].

Measures of performance are assessed by observing an operator's tasks skills, strategies, or procedures used to accomplish tasks. Here, quantitative metrics may also compare task execution time between members, human and robot idle time, or human-robot distance [137]. In particular, the motion has been observed through concurrent measures of average separation distance [138].

Measures of effectiveness observes the 'goodness' of the quality and execution of tasks while taking in dependencies into consideration that arise from the environment or luck. Response time and accuracy are key metrics here, but physiological measures or behavioral metrics through questionnaires are also telling [139]. A metrics of management of teams can describe how teams fan out, what intervention response times are, and how the level of autonomy impacts teams [100].

Individual Performance is another perspective for evaluating human-robot teams. These metrics range from exploring individual mission output to cognitive, internal measures of individuals. Some metrics measurable in humans and robots include situational awareness, workload influence, and mental vs. physical workload [140]. Social standards also come into play through interaction characteristics, persuasiveness, trust, engagement, and compliance [100]. Other measures of judgment, passive perception and judgment of motion can be useful measures of perception [100].

Human Performance often considers preferences and their fluency working with robotic counterparts [140]. Here, humans that are system operators may be assessed by their decision making, strategy, adaptability, mental model accuracy, and general workload [100].

Robot Performance , while often constrained by the limitations of the robot, still have interesting observable metrics. For example, Shah reviews metrics associated with physical interaction through response time, availability, proximity to physical interaction, and duration of physical interaction [133]. Robot performance has also been measured through self-awareness, human awareness, and autonomy, as well as through manipulation measures including the degree of mental computation and contact errors [100].

Another category of metrics is conceptual Measures of Teaming. The literature across several domains describes teamwork as being influenced by communication, collaboration, and coordination, among many other factors. However, the definition of these more conceptual terms is not agreed upon. The lack of definition and the breadth of these concepts results in difficulty in discerning an absolute measure for each. While these concepts are important to consider, not every team or scenario requires all three.

Communication is defined as the "expression of exchange of information between two (or more) parties" [50]. The format of communication can shift depending on the agents, their capabilities, and the necessary information to translate. Humans communicate through language, posture, facial movement, gesture, and more, while robots can convey information through auditory, visual, or physical means. Both convey a state, intent, or awareness, and to do so effectively requires perception, recognition, evaluation, and common ground. While humans have a larger capacity for transmitting or interpreting communication, robots require a better understanding of their communication needs.

For robots, specific metrics can be defined by the type of communication they are capable of. Auditory communication can be through speech (natural language) or sound (beeps, noise). Visually, they may communicate through text (words, keywords, phrases) or light (color, frequency). Physical communication is also possible through haptic feedback (vibration, physical), or motion (waving, turning, movement, gaze). This signaling can be any combination of the aforementioned methods which depends on whether or not signaling occurs human-human, human-robot, or robot-robot. Different teams have different needs for communication type and measuring different types of feedback can be simple (movement) or difficult (gaze and intent). Communication can be measured through quantitative and qualitative means. Metrics to define communication can have a large breadth by answering questions such as did they communicate the right information, how much did they communicate (bandwidth), was the communication efficient, and how successful was the communication effort (did the other party understand the information) [141, 142].

Coordination is defined as "the harmonious functioning of the group or ensuring that two or more people can work together properly" [50]. Commonly seen used to describe working closely, coordination implies working when tasks have unique sub-tasks. Coordination refers specifically to the way or method by which team members work together. It implies the organization, planning, and strategy of activities to function nominally without hiccups or breakdowns between teammates.

While the type of taskwork and external factors (randomness) can play a role in effecting coordination between team members, the team composition, training, and individual capabilities can just as much impact coordination measures. Coordination is key for developing shared goals and intentions between team members, in order to execute nominally.

Coordination can be measured by investigating physical interactions, scheduling, and task handoffs. The quality of physical coordination may be measured through the resources exchanged or placement of agents to one another as they work together. If agents work jointly together to hand off taskwork to one another, the speed, efficiency, and usefulness of this exchange can also be investigated. Organization and scheduling of agents and taskwork can also provide metrics to define coordination such as planning, foresight, and monitoring.

Collaboration is considered as a "joint activity involving two or more parties working collectively to achieve a common goal" [50]. This concept encompasses the partnership and collusion between team members socially to perform taskwork together. To do so, members also need to share knowledge, intention, and goals amongst each other. The mutual sharing of each member's perceptions and evaluation can also be expected it contributes to the larger goal. Collaboration is not how well the execution of tasks are completed (coordination), but rather the mutual kinship and participation between members to support the common goal and each other.

Metrics of collaboration are more qualitatively associated with the cognitive load, social teaming, and shared mental states. Making sure teammates are aligned on the goals of the mission and are aware of each other's mental states is key to understanding collaboration [23].

Observing the many aspects of human-robot teams highlights the many interesting dynamics of teaming and the corresponding metrics. While there does not exist a comprehensive list of metrics which effective human-robot teams should fulfill, prior research does demonstrate the breadth of metrics that have been proposed.

4.3 Expanding Work Allocation Measures to Human-Robot Team Metrics

Feigh, Pritchett & Young defined a series of categories for evaluating the allocation of work in teams. These metric categories define general measures of HAI from a work allocation perspective.

Workload/Taskload metrics defines task type, taskwork breakdown and definition, and taskload on the team and experiences by individual team members.

Mismatch in Authority/Responsibility notes when there is a difference between which team member is assigned to execute a task and which team member is responsible for its outcome. This may require teamwork actions by the responsible teammate such as monitoring or confirmation.

Stability of Human's Work Environment defines a measure of un/predictability in the environment or which the team will need to accommodate.

Coherency of the Allocation metrics examine for obvious effects that group (or break up) each agent's work in a manner that cannot be sensibly abstracted, such that incoherent allocations require excessive coordination due to, for example, giving different agents activities that are tightly coupled or that need to use the same resources. These metrics can be derived from values that capture efficient work practices, resource conflicts and coordination, and the distribution of functions of work on different agents.

Interruptions metrics note the profound disruption that interruptions can have on an individual's performance, whether they are completely unexpected or whether they reflect events that may be generally expected to occur at some time. Either way, an interruption is defined and counted as an event where one team members interrupt another, i.e. requires the other to stop a current activity and pay attention to communication and, potentially, perform another action.

Automation Boundary Conditions metrics note the extent to which machines are typically designed to only operate effectively within some given environmental boundary conditions. These metrics note the degree to which machine agents are placed close to (or beyond) their boundary conditions by a given work allocation, and the impact this will have on mission performance.

System Cost + Performance metrics reflect the overall implementation and operating costs of the team and the predicted performance of the team relative to mission goals. These metrics should be measured in all conditions and scenarios in which the team may operate, including both ideal, nominal conditions and foreseeable off-nominal conditions that the team is expected to accommodate.

Human's Ability to Adapt metrics reflect the value of team designs that allow their human team members to adapt their behavior to context. For example, metrics can evaluate the degree to which human teammates can adopt any of strategic, tactical or opportunistic cognitive control modes in response to different available times to execute their actions and different task demands.

The human-robot teaming metrics and work allocation metrics overlap on several key points. While we cannot suggest a definitive set of metrics to measure human-robot teams, we can show how HRI metrics can be thoughtfully measured from a work allocation perspective. Table 4.1 breaks down the eight work allocations metrics and how the human-robot teaming metrics fall into these buckets, Table 4.1.

Function Allocation Metrics	Robotics/HRI		
	Measures of Mission Performance	Measures of Teamwork	Measures of Individuals
Workload/Taskload taskwork breakdown and definition	Task Type Taskload metrics Mission Metrics (duration, execution) Taskload	Teamwork Actions Workload influence Action Dependency	
Mismatch in Authority/Responsibility differences between agent assigned to and responsible for outcome of single task		Physical/Cognitive Interaction Communication Management Teamwork actions Collaboration	Roles (within team) Trust
Stability of Human's Work Env. how humans react to unpredictable/dynamic changes	Mission success	Adaptability Shared mental models Spatial proximity (distance between members)	Situation awareness Communication Observability Perception Predictability
Coherency of FA roles, efficient work practicies, resource conflicts and coordination, and work assignment		Coordination Measures of sub teams	
Interruptions interruption quantity + quality		Idle time Communication/Information exchange	Cognitive interaction
Automation Boundary Conditions success of automation behavior in and out of boundary conditions	Performance metrics	Adaptability	Autonomy Self awareness
System Cost + Performance safety/robustness in off-nominal scenarios	Total failure time Failure recovery Operating Costs Collective Performance	Communication Coordination Physical Interaction Social measures	Trust Shared mental model Response time
Human's Ability to Adapt cognitive control modes, strategy (tactical, opportunistic), scheduling		Interdependencies Mental interaction Shared mental model (Group) Strategy	Fluency (Individual) Strategy Adaptability

Figure 4.1: Overlap between work allocation and human-robot teaming metrics

Two clear areas have shared metrics and standout as areas that have been well studied in both areas: mission performance and team interaction. These are two areas this work will explore in expanding work allocation measures for human-robot teaming through comparisons within multiple teams: single team performance and interaction, and individual performance.

4.4 Operationalized Metrics for Computational Analysis of Work Allocation in Human-Robot Teams

Using these definitions and the prior modeling constructs, we can define a group of work allocation metrics that will measure the range of teaming constructs identified in the literature as being important. Different interactions, team composition, and taskwork will all impact different aspects of teaming and require different types of analysis to understand and improve specific aspects of the team and teamwork. While human-robot teaming has a multitude of important metrics and perspectives to consider, we can categorize the list of teamwork metrics necessary for work allocation to the following two areas: overall mission performance and teamwork. In addition, while we consider many factors, this work, in particular, does not consider social robotics and measures.

Main Mission Metrics

Performance Metrics represent metrics across all team members and different work allocations. We will observe the aggregate values for each mission. Typical mission performance metrics include total mission time, total idle time, total failure time, and total taskwork time.

Taskload Metrics are metrics that breakdown taskwork further into detail per agent, work allocation, and cross work allocations. Within taskload, we will look at total taskload and the ratios of other task types such as regular task actions, teamwork actions, failure actions, traversal action, and fetching actions. Teamwork actions can be further divided into groups of confirmation, control, command, monitoring, and joint actions. These metrics will provide an indication of the mental and physical workload on the agents within a mission.

Teamwork (Interaction) Metrics

Coherency within teaming is the distribution of work between agents or sub-teams and their awareness of each other's workload. Sub-teams may arise from working closely together or recovering from failures in the mission. Different sub-team compositions will have different work distributions. One measure of coherency is how often and how long these pairs or sub-teams interact with each other, and the total workload of these sub-teams. More specifically, by observing the sub-teams that form through their interactions with the GIAPE resources, which will differ for each agent given their role, we can see what common sub-teams form and observe the cognitive impact that these work allocations have on individuals. We can also apply performance measures for each sub-team and look at the total information and physical transfer that may occur. In doing so, we may be able to understand which sub-teams are most successful.

Spatial Proximity metrics will investigate the interaction between agents working jointly, closely, and remotely in different teams within Function Groups (logical groupings of taskwork) across different work allocations.

Interdependencies will investigate which actions within the taskwork are independent or dependent on one another. In particular, by observing how the team works through the mission, we can identify joint, close, and remote work.

Conceptual Measures of Teaming here will represent simplified measurements of these terms. Due to the nature of these components, our evaluation of them are not qualitative values, but rather comparisons of how the requirements of these may change by each mission's performance.

- Collaboration as we've defined requires the cognitive alignment of team members and we will measure this through how the agents set or get the interaction resources GIAPE (goals, intent, actions, perception, evaluations) and the other types of resources (information and physical) in the mission. For example, when two or more agents must share a resource, we assume some requirement of collaboration between the two. This effect is magnified when we observe cognitive resources (GIAPE).

- Communication is simplified as an information exchange between agents and will consist of measuring which pairs or sub-teams communicate with each other most often. This is specifically measured through information resources that represent knowledge that is being shared between two or more agents. This exchange is captured when one agent sets an information resource that is later retrieved, or gotten, by another agent.
- Coordination metrics are more specifically the interaction of agents physically with shared resources and the amount and frequency of teamwork actions. When agents share the use of a physical resource (a screwdriver, screws) and work closely or event jointly together, we can assume the requirement of coordination between the pair. For teamwork actions, we evaluate the amount of cognitive coordination to complete these actions through the pairings which are usually between responsible and authoritative agents. This includes confirmation between members for correct action completion, monitoring of robots humans may be responsible for, control or commanding of robotic agents. By investigating these teamwork actions, we may learn which groups of actions may require the most coordination or the least.

4.5 Closing Thoughts

Future human-robot teams will require improvements to the existing portfolio of performance metrics. As the interaction grows increasingly complex, metrics need to expand beyond robot performance to include measures of cognition, team health, and relationships.

Here, we've expanded upon the traditional metrics found in HRI literature by utilizing cognitive engineering methods to analyze work allocation. The result is a broad set of metrics for analyzing human-robot teams that draw upon a comprehensive framework of human-robot teaming. These metrics analyze human-robot teaming from a mission, interaction, and teamwork perspective. This work describes the multitude of metrics by which human-robot teams can be evaluated, as well as the usefulness of work allocation as a method for team design.

The following chapters will implement the theories discussed thus far of human-robot teaming components and metrics with empirical work and analysis through case studies that examine computational simulation results, a HITL study of a simulated team, and the verification and validation of this work.

CHAPTER 5
ANALYZING HUMAN-ROBOT TEAMS AND INTERACTION

While there are many methods by which to analyze human-robot teams, we have seen in Chapter 2 the flaws and advantages between too-specific and too-broad practices. This chapter will dive further into work allocation combined with simulation and modeling to provide a full analysis of the human-robot team in questions. This analysis provides the capability of breaking down teams by cross-comparisons, cross work allocation, and cross agent evaluations via various visual graphics.

5.1 Work Allocation As a Methodology for Human-Robot Teaming

As noted earlier, work allocation is the team design of assigning different tasks to different agents [36, 33]. Designers often start by identifying the taskwork required for the team to complete its mission, either through Hierarchical Task Analysis or Work Domain Analysis [44, 47]. These analyses may identify a single sequence of action, or they may identify feasible sets of actions as important to the mission or team flexibility in how the mission goals may be achieved, and how variable the task environment is.

Previous research in allocating work focused on performance metrics such as total time to complete tasks, human involvement time, wait or idle time, resource consumption, or task-specific performance measurements [37, 38]. Other methods focused on estimating workload, based on the number and type of task that the human is performing, which some have proposed trading off against performance estimates to find the optimal function allocation [39]. Other work addresses the reliability of work allocations, quantifying and minimizing the probability of failure and human error [40]. Finally, there is a large body of research in aviation and related domains that focus on human performance metrics such as situation awareness and skill degradation associated with different function allocations

[41].

Pritchett, Feigh, & Young [36] use modeling and simulation as a method to evaluate the allocation of work in teams. With this method, designers can quickly evaluate different potential team designs spanning multiple teaming combinations, work allocations, task timing and more [143, 144, 145].

I have already developed computational simulations to evaluate work allocations for human-robot teams in spaceflight missions. These case studies explored a range of work allocations within human-robot teams in simulations of many relevant components of work: robot failures, recovery and response to failures; teamwork actions in general, with specific models of HRI; communication or information transfer, and sharing of physical resources such as tools [118, 51, 52]. Metrics have been expanded beyond typical measures of mission performance to also include metrics that capture relationships between team members.

Through the expansion of work allocation metrics in Chapter 4, we can now attempt to implement these metrics in simulated human-robot teams. In this chapter, I will walk through a series of steps to investigate human-robot teams through work allocation and simulation. Part of the ongoing research effort involves developing a methodology to investigate effective human-robot teaming through work allocation. In doing so, we've developed a methodology that does the following: "(1) perform analysis of constraints that are inherent to the work, the work environment and the agents, (2) explicitly account for the teamwork inherent to human-robot teams, and (3) dynamically evaluate and identify the emergent effects within the teams' collective capabilities " [**Submitted: ijtsma2019jcedm**].

The resulting methodology was created with these requirements in mind, including the following:

1. Pre-simulation, Work and Competency Analysis: this step provides static analysis of the work to be performed by a group of agents in this case a human-robot team. This step also evaluates the capabilities of the agents, objectives of the scenario, dependencies in the work, constraints, and potential interaction points. Much of the

work here begins with a CWA or HTA paradigm.

2. Simulation, Identifying Emergent Behaviors within the Team: in this step, we begin a dynamic analysis through WMC and identify any emergent interaction behavior and interdependencies. The computation simulation of the team acts as a first glance into the effects of work allocation and action timing on the effectiveness of the team.

3. Post-Simulation Analysis: this step takes the raw data following the simulation and generates specific metric results across team compositions and allocations of work between agents. The breakdown of data from a high and low level of various components of the team and individuals provides aa detailed investigation on any impactful inputs to the team performance.

5.2 Defining the Work and Human-Robot Team Case Study

The previous two chapters have described the constructs associated with the conceptual model we believe are necessary to capture differences in human-robot teaming and our operationalization of that conceptual model into a concrete computational framework including specific metrics to evaluate different interaction and teaming configurations. Central to the modeling was our effort to fully capture the interaction between human and robotic agents including not just the course grained allocation of functions between team members but also the nuanced ways in which the team members perform joint and interdependent work.

To demonstrate the merit and use of the metrics, a case study was analyzed. This case study involves an extra-vehicular mission in outer space, in which humans need to work together with various robots to maintain the spacecraft. The tasks to be conducted for the maintenance mission are the inspection of the space vehicle's exterior at three locations. At each location, the condition of the exterior panel needs to be checked, as well as the wiring and filter underneath the panel. If it turns out the condition of these components has deteriorated, they need to be replaced or repaired. The scenario was designed so that each location needs a single repair or replace task (in order, panel replacement, filter replacement and repair of wiring), see Figure 5.1.

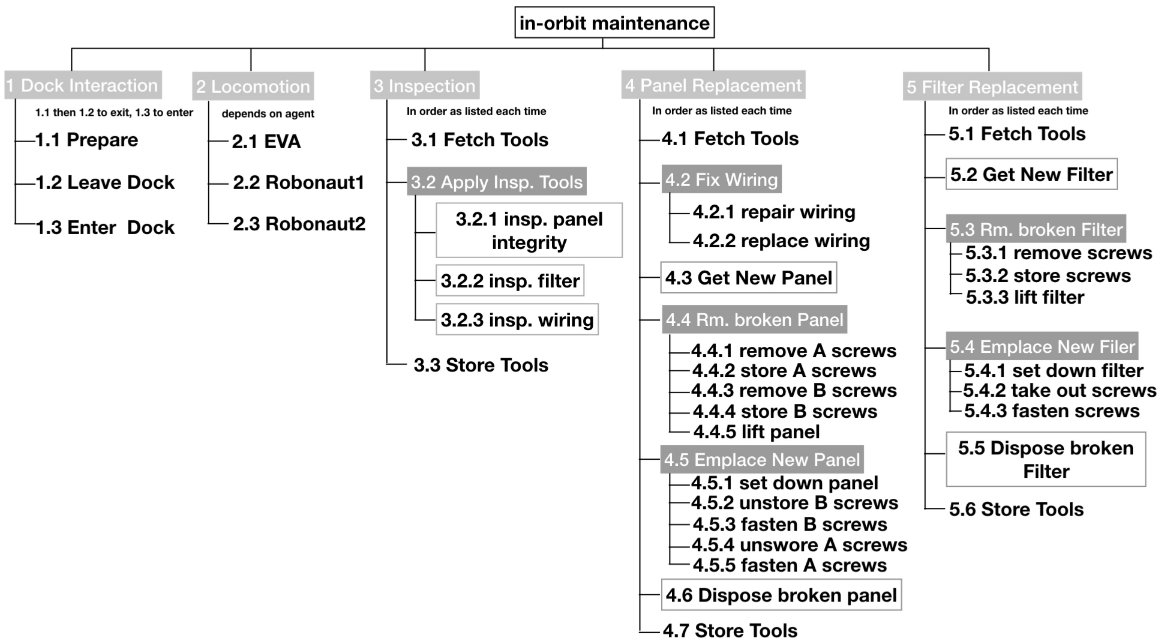


Figure 5.1: A hierarchical task analysis of the in-orbit maintenance mission.

The computational simulation requires as its input 1) a description of the work using one of several options. Here we chose a hierarchical task decomposition, 2) a definition of team composition, i.e. the role an agent takes on and their responsibilities and 3) the allocations of work, or work allocation, to different agents. Together, these items determine the type taskwork agents will complete, its physical proximity to other concurrent tasks (joint, close, remote) and any teamwork actions that are associated with the executing and responsible agent pairs. Any of these three inputs can be altered to produce different possible combinations of team interaction. In this work, we create and examine only four alternatives.

5.2.1 Team Composition

The agents in the scenario include several of the robots that NASA is currently using or developing: a humanoid robot that can operate outside the space vehicle (e.g., Robonaut), a free-flying robot that can perform simple inspection tasks (e.g., mini-AERCam) and a Remote Manipulator System (RMS) to do heavy lifting (e.g., Canadarm). Furthermore, human agents include an extra-vehicular (outside the vehicle) astronaut, an intra-vehicular (inside the vehicle) astronaut and the Mission Control Center (MCC) on Earth.

The ability of each robotic or human agent to perform each action was formally defined in four levels, which are based on Johnson's definition (Johnson et al., 2014):

1. The agent can independently perform the action.
2. The agent can perform the activity, but support from another agent can improve reliability.
3. The agent cannot perform the action on its own but can contribute.
4. The agent cannot perform the action.

When an agent is at Level 2-4 for these capabilities, the agent can be supported by another agent through joint activity, command or control inputs and/or through monitoring and confirmation actions. All robotic agents were assigned the role of robotic peers and were assumed to require input from a human agent through command or control, and verification of their actions through real-time monitoring or confirmation. The human agents were assigned the role of supervisor, thus being able to provide control inputs to and verify their robotic peers.

Based on this definition of capabilities, two possible team configurations were identified. The first team configuration consists of the EV astronaut, the IV astronaut, and the humanoid robot. The second team configuration consists of again the EV and IV astronauts, with the remote manipulator system and the free-flying robot.

Table 5.1: Two allocations of work for team configuration 1.

Function	Team 1 Allocation 1			Team 1 Allocation 2		
	EV	HR	IV	EV	HR	IV
Inspection	✓			✓	✓	
Panel Replacement	✓	✓			✓	
Filter Replacement	✓	✓		✓		
Wiring repair	✓			✓		
Fetching		✓		✓		
Controlling and monitoring Humanoid			✓(MC)			✓(MC)

Note: EV = extra-vehicular astronaut, HR = humanoid robot, IV = intravehicular astronaut; MC = manual control with real-time monitoring, CS = command sequencing with confirmation.

Table 5.2: Two allocations of work for team configuration 2.

Function	Team 2 Allocation 1				Team 2 Allocation 2			
	EV	RMS	FFR	IV	EV	RMS	FFR	IV
Inspection			✓				✓	
Panel Replacement	✓	✓			✓	✓		
Filter Replacement	✓	✓			✓	✓		
Wiring repair	✓				✓			
Fetching			✓				✓	
Responsible for RMS				✓(MC)				✓(MC)
Responsible for FFR				✓(CS)	✓(MC)			

Note: EV = extra-vehicular astronaut, RMS = remote manipulator system, FFR = free-flying robot, IV = intra-vehicular astronaut; MC = manual control with real-time monitoring, CS = command sequencing with confirmation.

The team configurations and work allocations (referred to as FA) are shown in Table 6.1 for the first team configuration, and Table 5.2 for the second team configuration.

5.2.2 Allocations of Work

Based on this definition of capabilities, we target three possible ways to allocate the work or work allocations were identified. They were constructed such that all actions could be performed with at least level two of the list above, either by individual agents or a combination of two or three agents. When two or three agents were jointly performing an action, the teamwork modes were specified as part of the work allocation. For example, as a robot was assigned an action, but a human was operating and verifying the robot, a control mode (manual versus commanding) and verification mode (monitoring versus confirmation) were specified.

Team 1

Work Allocation 1 (FA1) The EV astronaut performs the inspection of panels, filters, and wiring and performs the replacements and repairs. The humanoid robot assists with any heavy lifting tasks (the EV and humanoid jointly perform any lifting tasks during a panel and filter replacement) and fetches the required resources. The IV astronaut is responsible for the robotic operations and controls the robot through manual control inputs.

Work Allocation 2 (FA2) Compared to FA1, the humanoid and EV astronaut is working further apart from each other. The humanoid performs all actions related to panels (inspection and replacement), and the EV astronaut all actions related to the filters and wiring (inspection and replacement). The actions have purposely been allocated to individual agents only, except for inspection, and it is hypothesized that this will result in fewer dependencies between each agent's activities. The IV astronaut still controls the humanoid robot.

Team 2

Work Allocation 1 (FA1) The free-flying robot performs an inspection of panels, filters, and wiring and fetches any required physical resources. The EV astronaut performs the replacement of panels and filters and the repair of the wiring. The RMS supports any of the heavy-lifting tasks, but in contrast to Team 1's FA1, it does not need joint support from the EV astronaut for these actions. The IV astronaut is responsible for both robot's operations and controls the RMS through manual control inputs and the free-flying robot through command sequencing with confirmation following each of the robot's actions.

Work Allocation 2 (FA2) This work allocation has the same distribution of the taskwork as FA1, but differs in who is responsible for and controls the free-flying robot: here it is the EV astronaut that is manually controlling the robot's actions. It is hypothesized that this will have a major effect on the time and teamwork metrics as it will significantly impact the dynamics of the work of the EV and IV astronauts.

5.3 Analysis Kit Results: Investigating Work Allocation and Team Composition

First, I will investigate each team composition separately by comparing the two FAs for each team using the metrics mentioned in Section 3. We begin by discussing the overall quantitative mission metrics before diving into more detail on the teamwork and interaction metrics. Next, we will investigate a cross-team analysis of the four work allocations. These results will demonstrate how different team compositions can impact performance. It is important to note some of the figures mentioned are included in the Appendix section.

5.3.1 Analysis Kit Development

The metrics described alone require some scoping or context to provide thoughtful reasoning about these human-robot teams. The implications of metrics are valuable for assessing different work allocation decisions, mission and taskwork design, and teamwork within teams. I organized the analysis of our teams into three levels of abstraction by agent level, work allocation level, and cross work allocation level.

Agent Level Analysis at this level can be useful for investigating particularly impacted or un-impacted agents. Agent reports output quantitative metrics for individual mission performance, communication, physical interaction, and teamwork actions. This informs designers of an agent's total taskwork, how their time was split between busy vs. idle time, and how well they performed on their taskwork overall.

Work Allocation Level On a per work allocation level, we can begin to compare agent performance to one another and discover sub-teams that commonly work together. Specifically, it is easy to see the division of labor between each agent and that impact on the mission. This level combines the output of quantitative metrics for overall mission performance with taskload and action breakdown figures, idle and busy time for the team overall, and graph networks of communication and physical interaction throughout the mission. Additionally, the coherency and distribution of workload combined with the interdependencies of actions are apparent by comparing agents to one another. In particular, we can also observe the interdependencies of actions to one another by observing the frequency of interaction with resources as well as the impact agent roles have on their cognitive exchanges of goals, intent, perception, and evaluation.

Cross Work Allocation Level At this level, we can compare how different work allocations of the agents in a team can perform the same work. This comparison helps answer the question of whether there are allocations of work that have properties that make teams successful. Metric reports here compare mission time, taskwork and teamwork action breakdown, and cumulative taskload. In addition, we can also compare how a specific agent's performance and workload balance changes between work allocations. More importantly, by cross-comparing work allocations, we can observe what small changes to the team member's responsibilities and taskwork can impact conceptual measures of teaming and how the needs of the teams will differ cross work allocation.

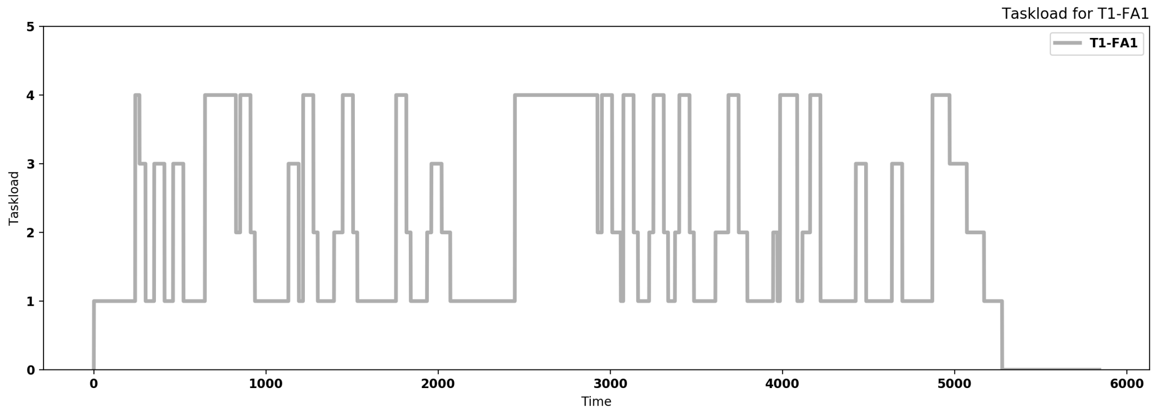
5.3.2 Team 1 Discussion and Analysis

Team 1 consisted of just three agents: an EV astronaut, an IV astronaut, and Humanoid robotic agent. By directly comparing the differences between each work allocation for this team (T1-FA1 and T1-FA2), we observe some interesting distinctions between these two work allocations.

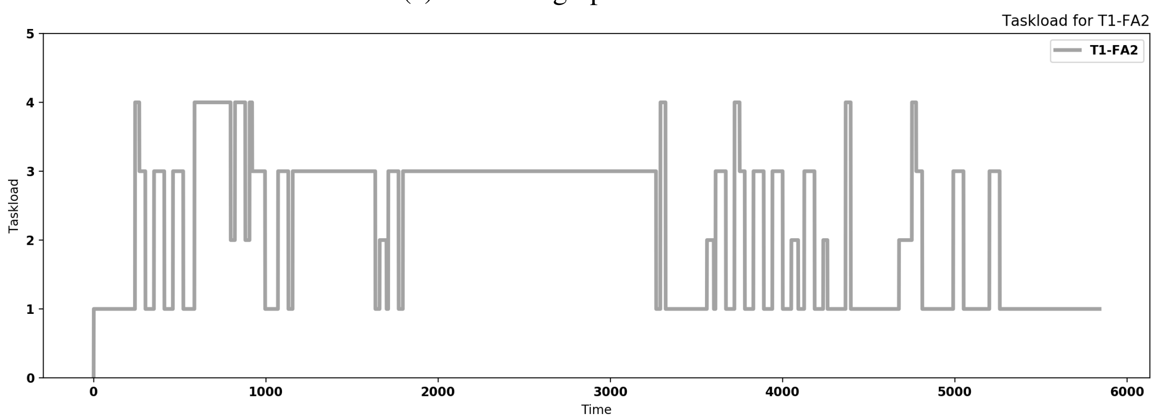
The first allocation (T1-FA1) had a shorter total mission, total task duration time, total busy time, and total idle time, see Table 5.3. This implies that in a shorter mission duration, the first team work allocation was consistently completing work without as much downtime. The first work allocation also has a lower total taskload and fewer teamwork actions, which support the difference in time. The second work allocation (T1-FA2) has higher results for all the aforementioned metrics. The highest of which is an increase in total idle time by 25.8% compared to FA1, see Table 5.3. For some teams, a higher idle time may indicate a much more ineffective work allocation. However, the fact that FA2 has more taskwork also suggests the increase in these metrics should not be surprising. When comparing taskwork for the team over time, FA2 has steady work for about 2000 seconds, whereas FA1 seems to vary in many peaks throughout the mission, see Figure 5.2a, 5.2b. Both work allocations peak at 4 tasks at any given time, however, FA1 appears to be working at that peak for many more instances compared to FA2. From these main mission metrics, Team 1's FA1 may appear to be a slightly better work allocation compared to FA2, see Table 5.3.

Table 5.3: Total mission metrics for Team 1 compared between the two work allocations, highest value in bold.

Total Mission Metrics	Team 1		
	WA 1	WA 2	Difference (%)
Total Mission Duration (s)	5276	**5841**	10.7
Total Task Duration (s)	11525	**12895**	11.9
Total Busy Time (s)	9645	**9745**	1.0
Total Idle Time (s)	6183	**7778**	25.8
Total Failure Time (s)	n/a	n/a	0
Total Taskload (count)	162	**169**	4.3
Total Teamwork Actions (count)	58	**71**	22.4
Total Physical Actions (count)	35	35	0



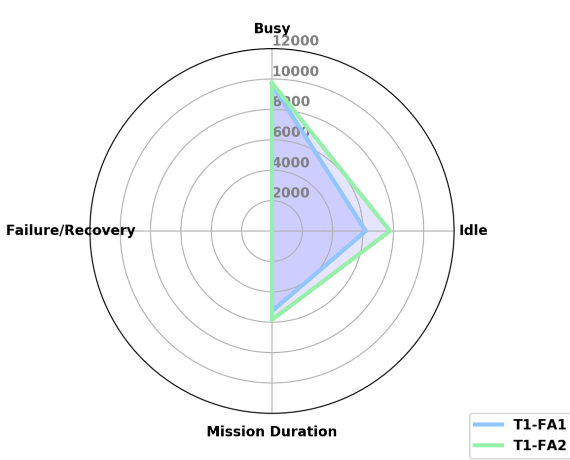
(a) Taskload graph for T1-FA1



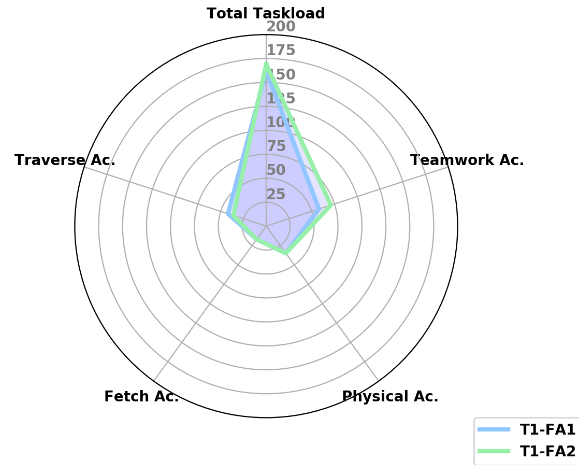
(b) Taskload graph for T1-FA2

Figure 5.2: Comparison of total taskload for Team 1's work allocations.

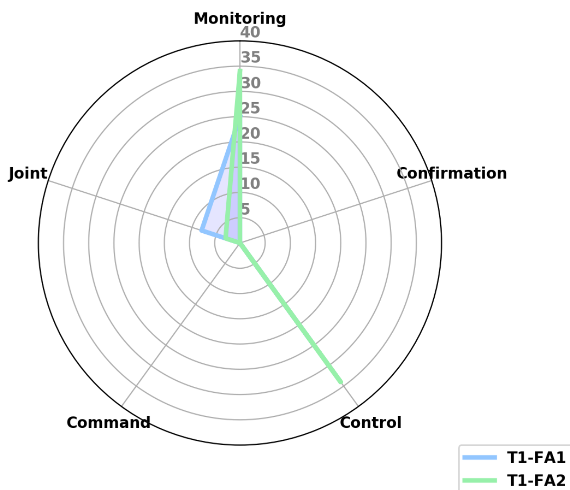
To further investigate these results, we can look more deeply into mission metrics and teamwork metrics. Figure 5.3a shows how the two work allocations performed with respective to busy time, mission duration, and idle time. Both teams have similar total busy times, however, FA1 has a significant decrease in idle time. Given that both work allocations appear to have similar total taskload, FA1 has more traversal actions and fewer teamwork actions than FA2, see Figure 5.3b. Interestingly for FA1, the focus of teamwork actions is within joint and monitoring actions. FA2, which had an increase in teamwork actions, required both the aforementioned teamwork actions and additional manual control actions, see Figure 5.3c. Finally, for communication and physical interaction, the first work allocation (T1-FA1) has less communication but more physical interaction than FA2, see Figure 5.3d.



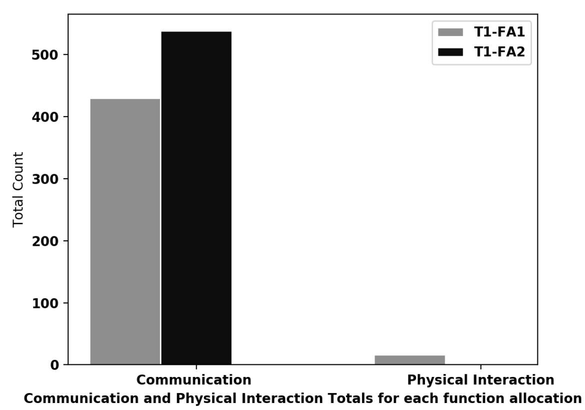
(a) Main Mission Metrics: Time spend busy, idle, failure recovery, and total mission duration



(b) Taskwork Action Metrics: Total taskload count and spread across different actions (ac).



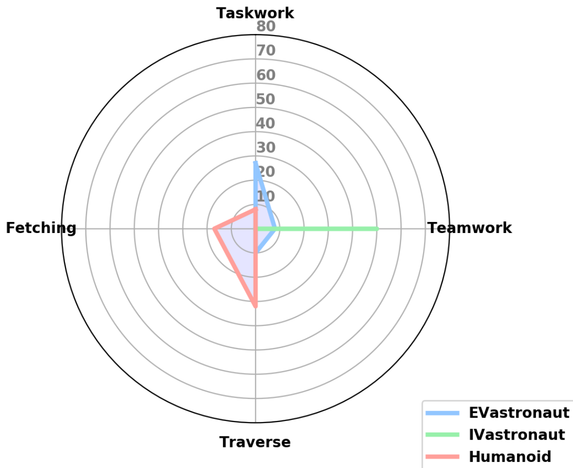
(c) Teamwork Action Metrics



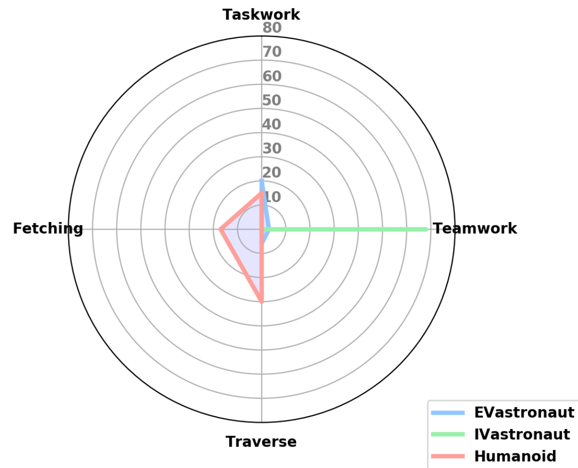
(d) Communication and Physical Interaction

Figure 5.3: Big picture overview comparing both FA for Team 1 for main mission and teamwork metrics.

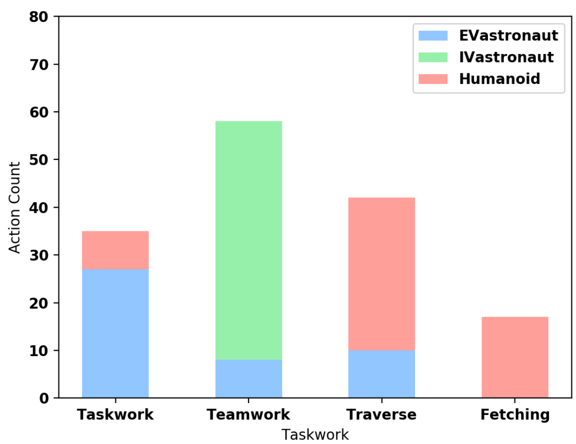
Although the first work allocation (T1-FA1) seems to accomplish the mission in a shorter time period, this performance comes at the cost of increased taskload on the human agents and management of physical interaction between members. To understand if this additional taskload is acceptable, we need to look at the impact on the agent level. We can observe comparing agent contribution to taskwork and teamwork, Figure 5.4, 5.5. For the first work allocation (T1-FA1), it is clear that EV astronaut has higher taskwork in their respective task types. In the second allocation (T1-FA2), the IV astronaut has an increased load in teamwork actions that reflects the Humanoid robot taking on some more regular taskwork, reflecting a shift in physical taskwork and traversal actions from the EV in the first work allocation (T1-FA1) to the Humanoid robot in the second work allocation (T1-FA2). This shift in workload causes changes the teamwork action load to be reduced overall for FA1. The transfer of work from EV astronaut to Humanoid increases the IV astronaut's effort expended on teamwork actions like control and monitoring, see Figure 5.5.



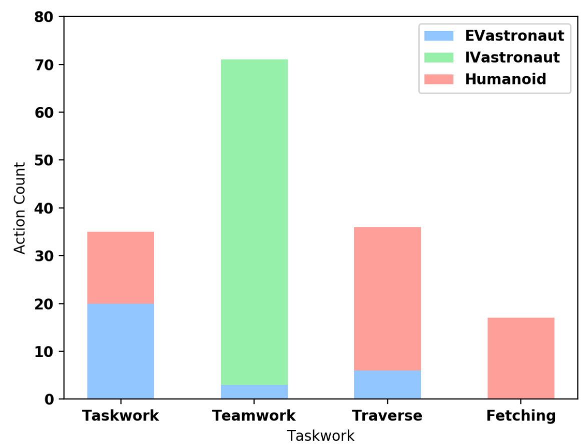
(a) Taskwork Breakdown between Agents in Team 1 FA1



(b) Taskwork Breakdown between Agents in Team 1 FA2

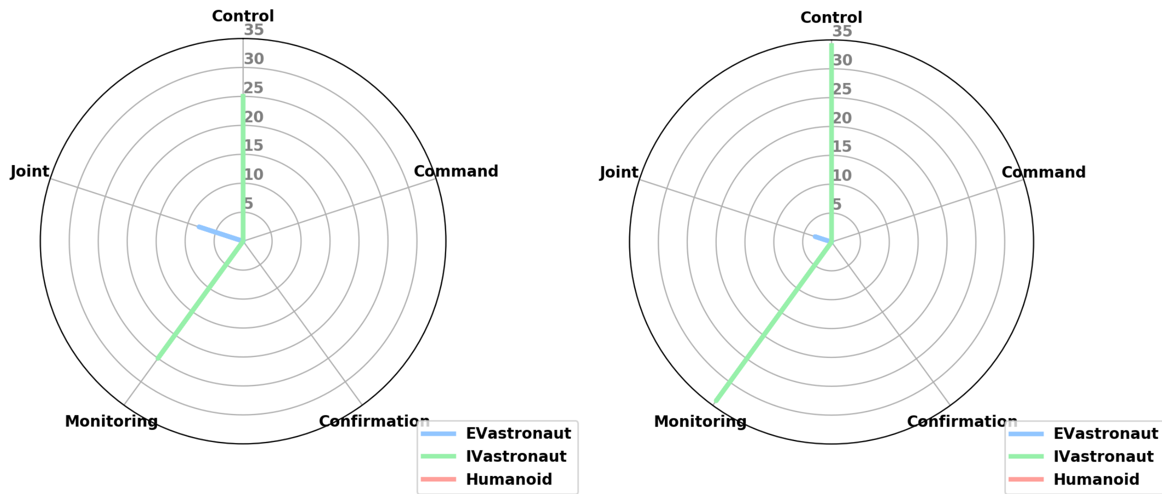


(c) Taskload spread to agents in Team 1 FA1



(d) Taskload spread to agents in Team 1 FA2

Figure 5.4: Spider and Bar Graph of Team 1 Agent Contribution to Taskwork

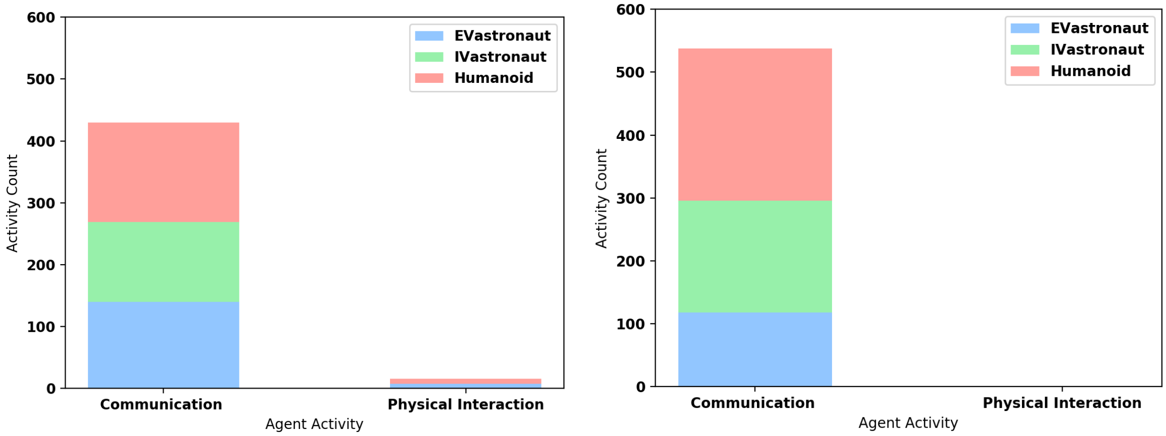


(a) Teamwork action spread to agents in Team 1 FA1

(b) Teamwork action spread to agents in Team 1 FA2

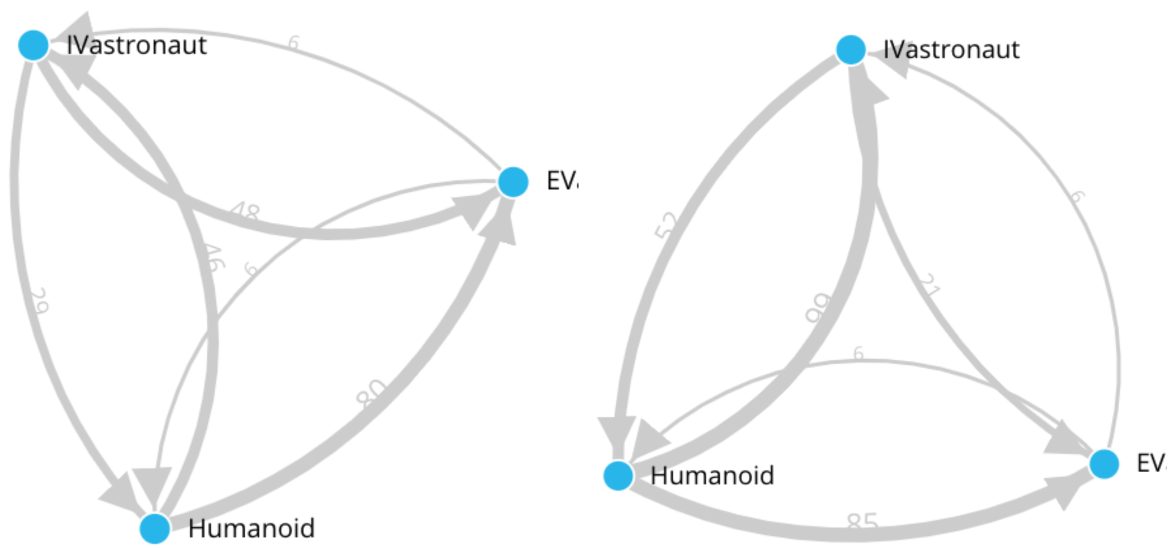
Figure 5.5: Team 1 Agent Contribution to Teamwork

Specifically, the IV astronaut also experiences an increase in monitoring and control actions overall from T1-FA1 to T1-FA2, see Figure 5.5. The IV astronaut performs more communication tasks in FA2 which mainly stems from the increase in direct communication between IV to both the Humanoid and EV astronaut, see Figure 5.7. Humanoid and IV communications become the top pairwise communicators in FA2, while Humanoid-EV, formally the highest in FA1, decreases slightly, see Figure 5.7. In a network of communication, the FA1 has a more evenly distributed load across the agents, although, in both work allocations, all agents communicate back and forth with each other. Physically, however, the instances of interaction between EV and the humanoid are eliminated in the second work allocation T1-FA2.



(a) Communication load to agents in Team 1 FA1 (b) Communication load to agents in Team 1 FA2

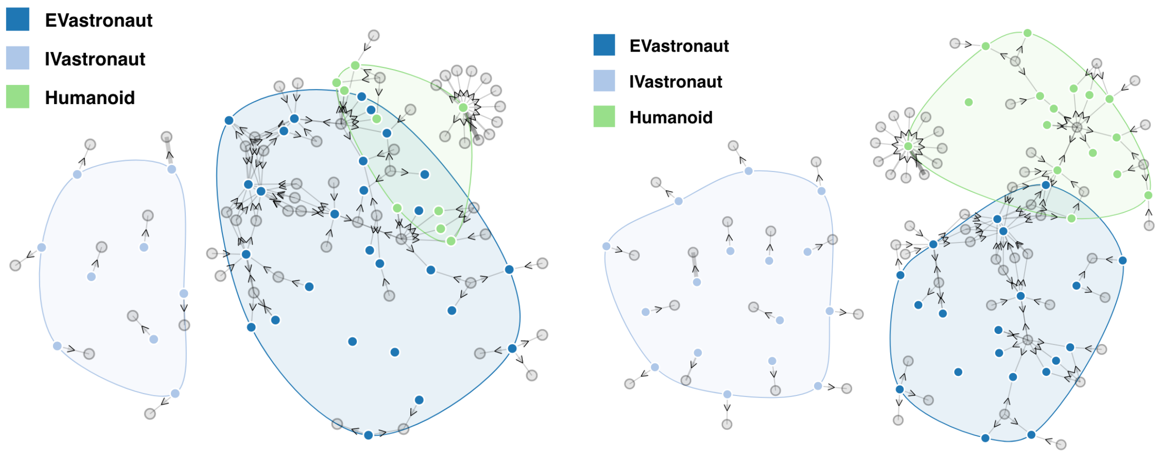
Figure 5.6: Team 1 Agent Contribution to communication and physical interaction



(a) Graph network of Communication in Team 1 FA1 (b) Graph network of communication in Team 1 FA2

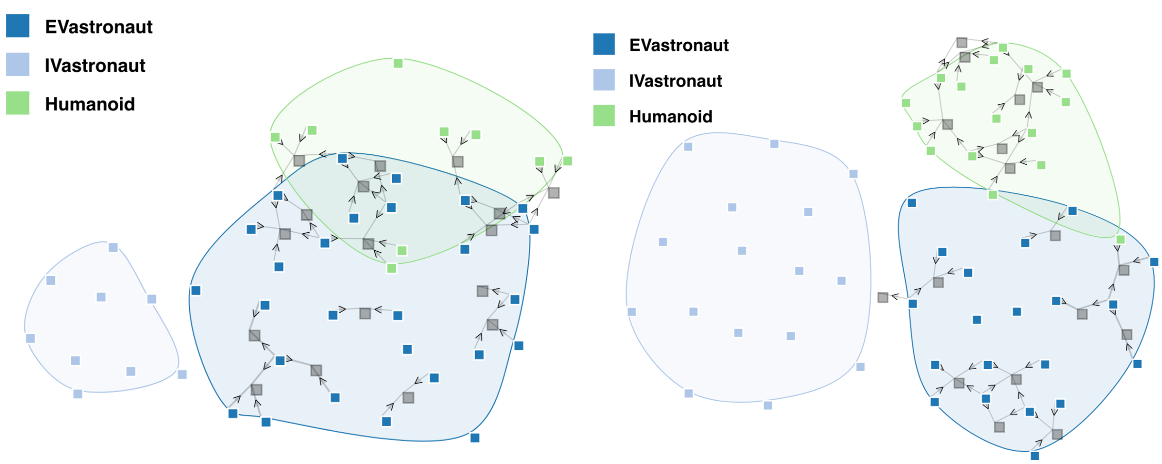
Figure 5.7: Team 1 graph network of communication

The physical interactions and communication between agents is a result of how interdependent the actions assigned to each agent are. An action is capable of *using* a physical resource or *getting* or *setting* an information resource. Therefore actions that are highly interdependent result in increases in these interactions. By creating a graph network of relationships between *actions* to *resources*, we can see the interdependencies of actions and agents to one another. Specifically, these networks have *action nodes* where the color corresponds to the agent that executed it and *resource nodes* which are gray. *Edges* between these action nodes and resource nodes depend on how the action interacts with that particular resource, via setting or getting information resources, (Figure 5.8), or using physical resources (Figure 5.9). Thicker line weight represents more instances of interaction with a specific resource.



(a) Team 1 FA1 (b) Team 1 FA2

Figure 5.8: Team 1 graph network comparison of action nodes (colored) and information resource nodes (gray) through amount of getting or setting.



(a) Team 1 FA1 (b) Team 1 FA2

Figure 5.9: Team 1 graph network comparison of action nodes (colored) and physical resource nodes (gray) through number of uses.

From these networks, we can see that the first work allocation had much higher coordination and collaboration requirements between the EV astronaut and the Humanoid robot. The clustering of the Humanoid action set (green) and EV astronaut action set (blue) are not separable due to actions sharing resources amongst each other. However, it appears from the second allocation that there is slightly less overlap in these shared resources. Additionally, while the IV astronaut's interaction with resources is separate in both work allocations, the number of resources interacted with for getting and setting and the total number of actions executed by IV are both higher in FA2, Figure 5.8.

Based on the roles of agents, we surmised a series of small network graphs that represent cognitive load and more conceptual measures of teaming, see Figure 5.10. IV serves as the Supervisor, all other human agents are Human Peers, and robots operate as Robot Peers. Each resource Goal, Intent, Perception, and Evaluation are set or gotten by the agents in the team. For team goals, in particular, we observe that IV with the role of Supervisor is in charge of setting and adjusting the goal throughout the mission. For collaboration and coordination on intent, perception, and evaluation, IV also takes on the most load in confirming the team's understandings of these concepts. Overall, the first work allocation has lower cognitive loads for IV and the Humanoid. However, EV experiences a reduction in cognitive load in FA2 all agents and aligning team intent is the most significant component amongst goals, perception, and evaluation.

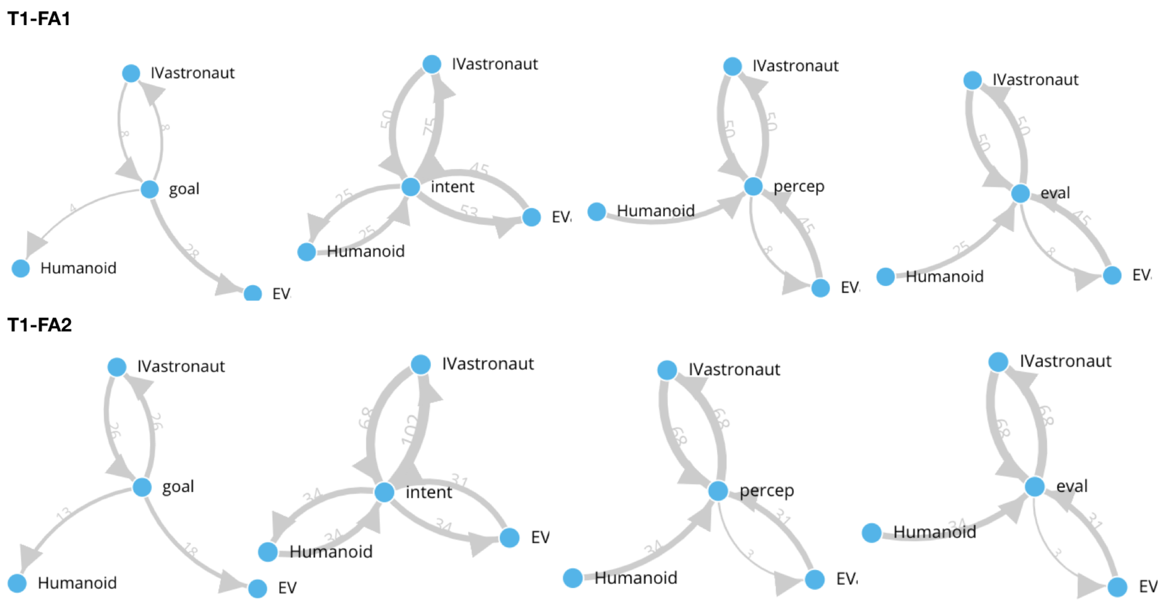


Figure 5.10: Capacity Resources for cognitive load comparison for both Team 1 FAs

Team 1 Conclusions

In Team 1, the first work allocation appears to have a lower mission time and lower idle time, which may seem more efficient than the second work allocation. FA1 also has a more even distribution of physical and cognitive load among the agents. However, FA2 has fewer action dependencies between agents, including entirely eliminating physical interaction. This can help reduce the required coordination planning for the teams. In FA1, EV and Humanoid agents are tied closely together due to the shared information and physical resources shared between them. In FA2, the overlap between action interdependency is much lower, which may be more desirable to reduce any potential bottlenecks. Depending on the goal of the team designer, higher or lower values for these assorted metrics can imply team success.

5.3.3 Team 2 Discussion and Analysis

Team 2 consists of more agents than team one, specifically operating without a humanoid robot and instead with a Free-flying robot and an RMS robot. In Table 5.4, we can observe the main mission metrics between the two work allocation for this team. The first allocation (T2-FA1) appears to perform better than the second FA(T2-FA2) for every metric with lower values except the total teamwork action count. T2-FA1 has a lower mission and idle time implying that this work allocation allowed the team to be more effective in completing the mission, with a reduction of 20.5% and 42.9% respectively. The first work allocation has a lower total busy time and less taskload than the second work allocation.

Table 5.4: Total Mission metrics for Team 2 work allocations, highest value in bold.

Total Mission Metrics	Team 2		
	Work Allocation 1	Work Allocation 2	Difference (%)
Total Mission Duration (s)	4109	**4951**	20.5
Total Task Duration (s)	8594	**12280**	42.9
Total Busy Time (s)	7695	**9520**	23.7
Total Idle Time (s)	8741	**10284**	17.6
Total Failure Time (s)	n/a	n/a	0
Total Taskload (count)	184	**185**	0.5
Total Teamwork Actions (count)	**81**	80	1.3
Total Physical Actions (count)	35, equal	35, equal	0

From the overall mission metrics, it appears that thus far the first work allocation performs better. By comparing the total taskload over time for both work allocations, both appear to have the same maximum of 6 actions simultaneously, see Figure 5.11. However, FA2 reaches this peak numerous times while FA1 only reaches this maximum once. The first work allocation (T2-FA1) has an overall higher taskload count at any point in time, while also completing the mission in a much shorter duration. Between the two, total task duration is the biggest difference of 42.9% for FA1 which has lower busy time, idle time, and total mission duration. This is interesting given the total taskload between the two work allocations differ by just one action, see Figure 5.12b. When observing the teamwork action breakdown, FA1 requires all teamwork actions (command, confirm, control, and monitoring). FA2 however, only requires monitoring and control, see Figure 5.12c. Physical interactions between the two work allocations do not change, while communication increases in the second allocation (T2-FA2), Figure 5.14.

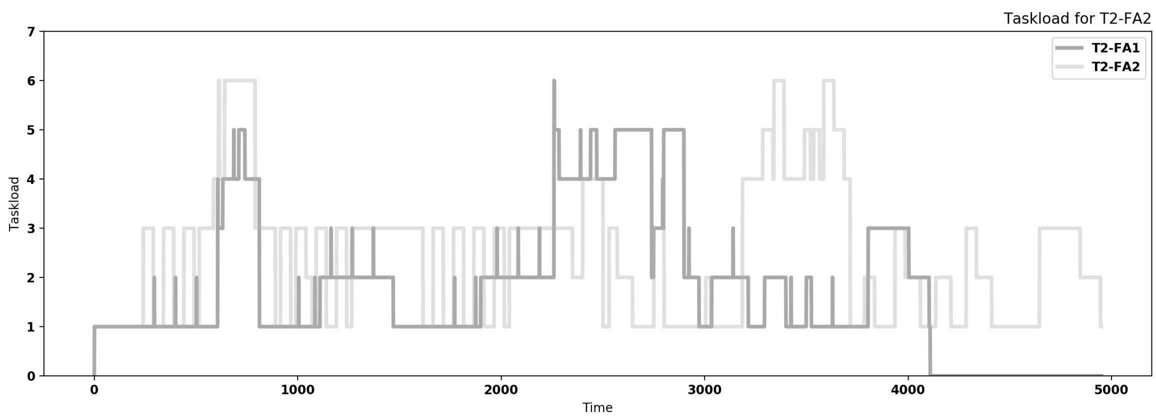
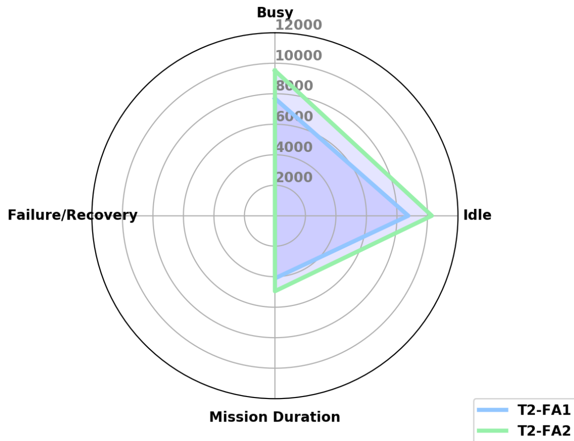
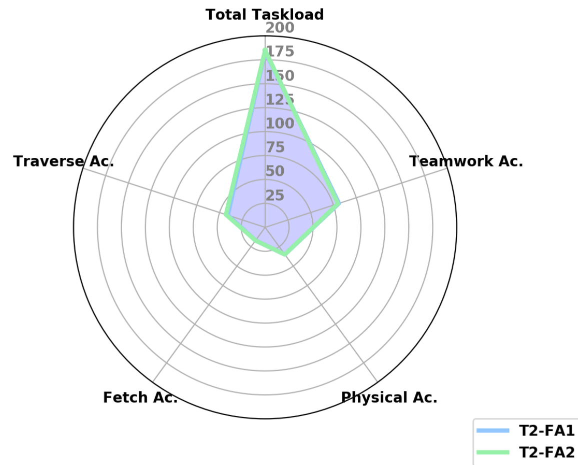


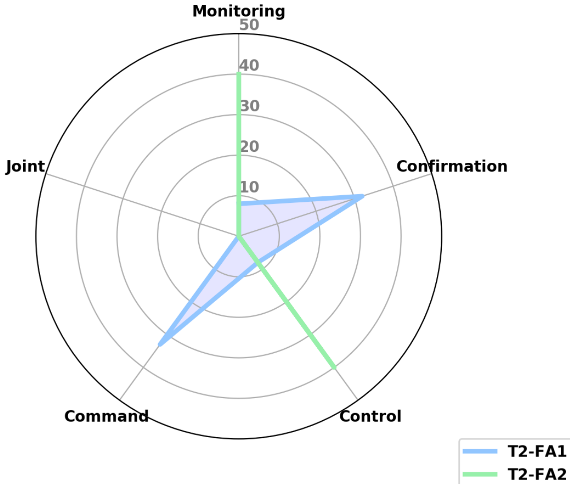
Figure 5.11: Team 2 comparison of total taskwork for both work allocations.



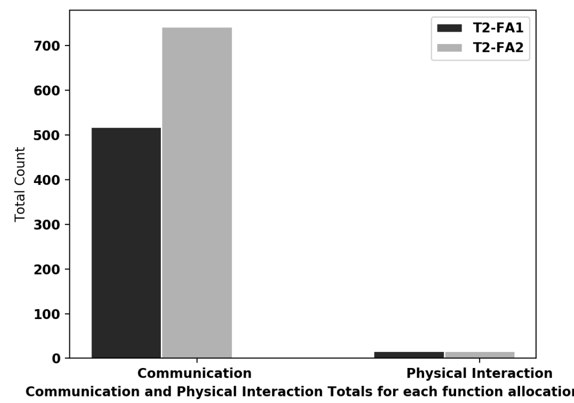
(a) Main Mission Metrics Breakdown



(b) Taskwork Action Metrics Breakdown



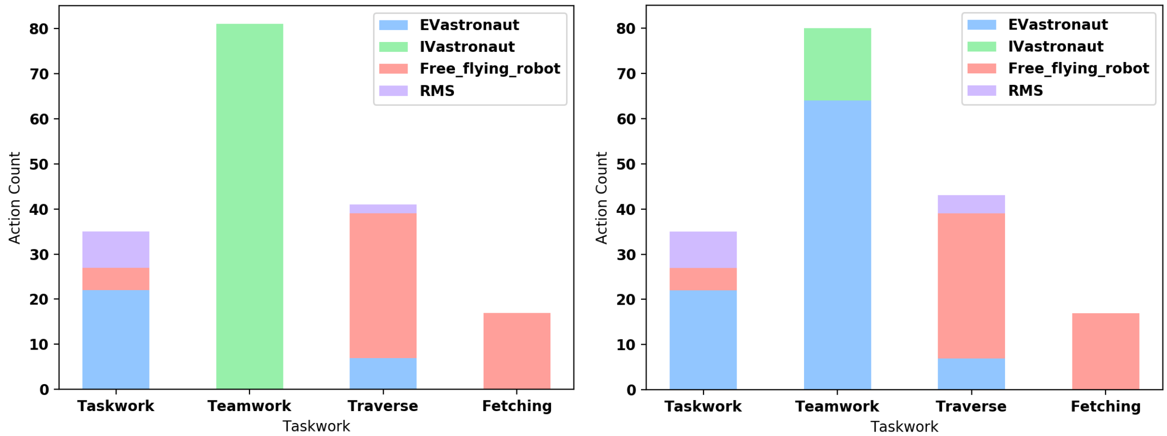
(c) Teamwork Action Metrics Breakdown



(d) Communication and Physical Interaction Breakdown

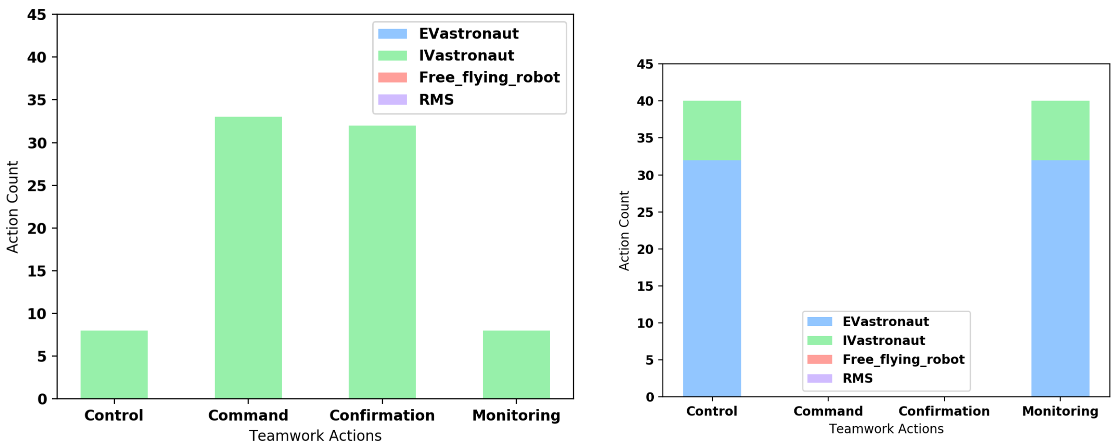
Figure 5.12: Team 2 Comparison of two FAs across mission and teamwork metrics.

While the main mission metrics tell us FA1 performs better, the difference in teamwork actions between the two work allocation requires more in-depth observation in individual agent performance. The Free-flying robot and RMS robot remain relatively unchanged in terms of taskload. EV astronaut shifts from completing physical taskwork and traversals to physical, teamwork, and traversal taskwork between T2-FA1 to T2-FA2, see Figure 5.13. Additionally, the teamwork contribution of the involved agents changes greatly from the first work allocation to the second, see Figure 5.13. IV performed all teamwork actions in FA1, consisting of some control and monitoring and many command and confirmation actions. In FA1, the teamwork actions become mostly performed by EV, requiring solely control and monitoring actions, see Figure 5.14.



(a) Taskwork allocation for agents in Team 2 FA1 (b) Taskwork allocation for agents in Team 2 FA2

Figure 5.13: Team 2 Agent Contribution to Taskwork

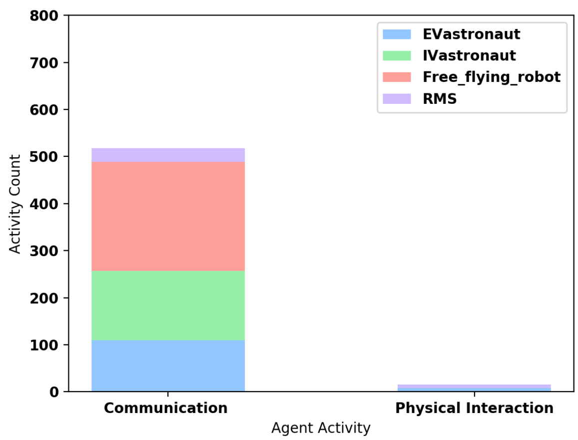


(a) Teamwork allocation for agents in Team 2FA1 (b) Teamwork allocation for agents in Team 2FA2

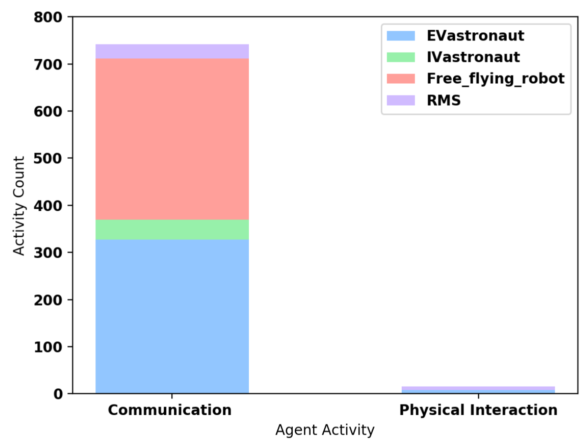
Figure 5.14: Team 2 Agent Contribution to Teamwork Actions

Regarding communication, IV astronaut's communication is reduced, while EV's and the Free-flying robot's increased. In the first allocation, the majority of the communication occurred between the IV astronaut and Free-flying robot. In the second, the most pairwise communication stems from the EV to the Free-flying robot, see Figure 5.16. The change in communication load to agents in most apparent in Figure 5.16, where line thickness represents load or weight. Here, a potential delay in communication can occur between

RMS and EV, as there is no direct line of communication between the two. Generally, the balance of communication in FA1 appears to be more evenly distributed compared to FA2, where EV has more communicative taskload.

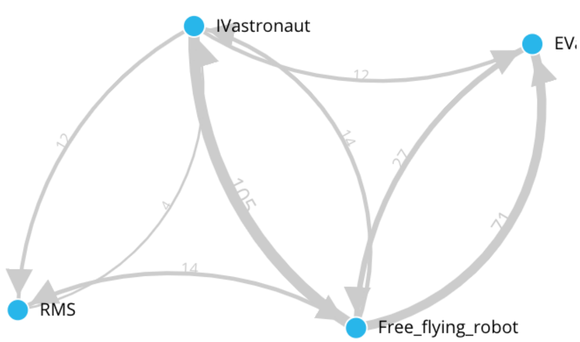


(a) Team 2 FA1

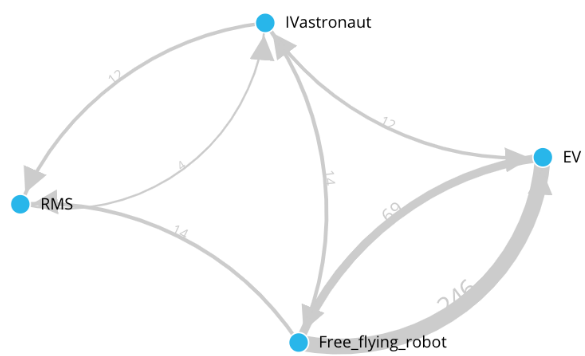


(b) Team 2 FA2

Figure 5.15: Team 2 Agent Contribution to communication and physical interaction



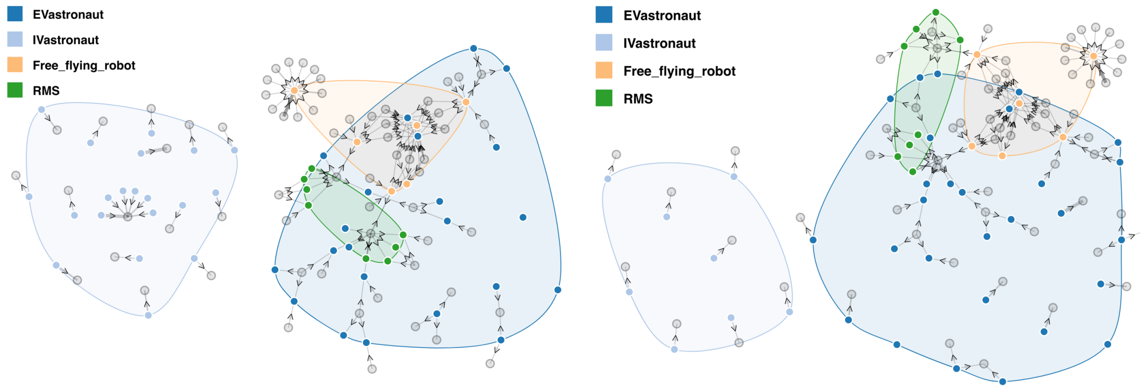
(a) Team 2 FA1



(b) Team 2 FA2

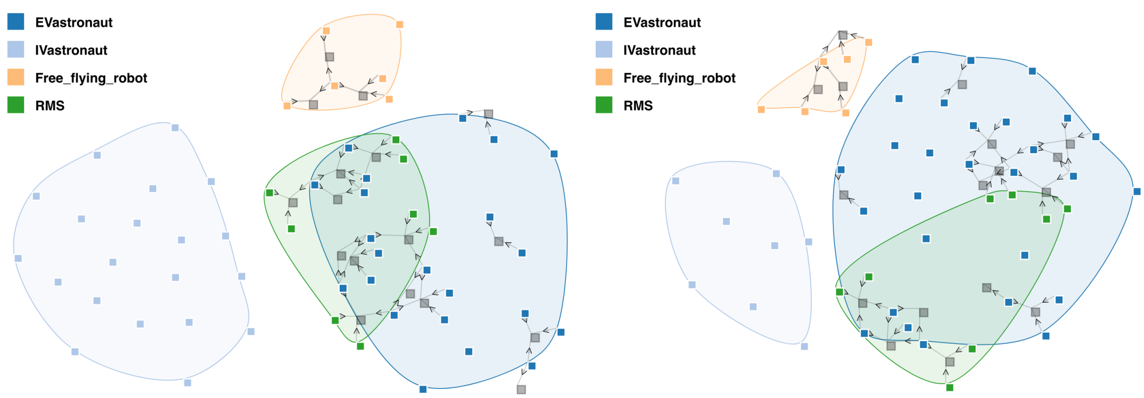
Figure 5.16: Team 2 graph network of communication

When observing action dependency for FA1, we see that EV, the Free-flying robot, and RMS have highly dependent actions sets, compared to IV who's actions do not interact with resources that other actions require, see Figure 5.18,5.17. FA2, in contrast, has fewer dependencies between the actions handles by agents, and IV also has a smaller set of actions. For information resources, the two robots and EV have action dependencies. Between the four agent action sets, EV and RMS have the most dependencies, which is most apparent in sharing physical resources.



(a) Team 2 FA1 (b) Team 2 FA2

Figure 5.17: Team 2 Action dependencies of agents comparison for shared information resources.



(a) Team 2 FA1 (b) Team 2 FA2

Figure 5.18: Team 2 Action dependencies of agents comparison for shared physical resources.

Observing the cognitive load for the first allocation, we see that IV astronaut is loaded with the most getting and setting of the cognitive resources, Figure 5.19. For FA2, the majority of the cognitive load is distributed more to EV. While IV agent as the mission supervisor is still responsible for setting the team goal's, EV works more closely with the other robot agents and thus becomes the main point of contact in aligning intent, perceptions, and evaluations throughout the mission.

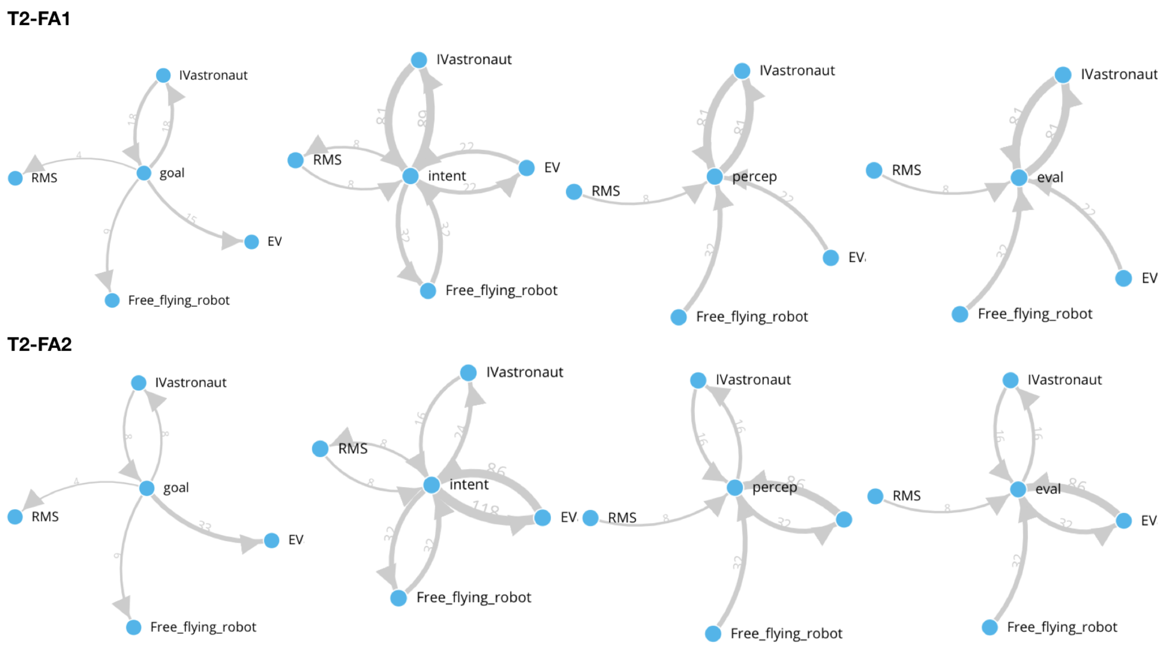


Figure 5.19: Capacity Resources for cognitive load comparison for both Team 2 FAs

Team 2 Conclusions

In team 2, we see that for mission metrics, the first allocation is better in terms of lower mission time, idle time, and more given a similar amount of taskwork completed. However, this comes at the cost of more teamwork taskload on the IV agent who is responsible for all teamwork actions. Additionally, in FA2, only two of the four teamwork actions in FA1 are required. While FA1 has an overall lower communication total, much of the load is shared with EV agent in FA2. FA2 also has fewer action dependencies between shared information and physical resources throughout the mission. For cognitive load, FA1 weighs heavily on the IV astronaut as the Supervisor to handle the entire team's coordination and collaboration to align mental models. Comparatively, FA2 spreads this load across IV and EV to create a more balanced cognitive load. The tradeoff, in this case, is high taskload and stress on one member for potential errors, data drops between communication and difficulty in transparency between MCC, IV, and EV.

5.3.4 Cross Team Analysis

To cross compare these four allocation teams together, we can run the same high-level analysis kit on both teams. Although the team composition across the two teams differs, the mission workload is the same. By comparing across all instances of teams, we can begin to look at how team composition and work allocation impacts performance.

Overall, for total mission metrics, we can point out that none of the four teams stands out as the absolute best. The strengths and weaknesses of these four teams is captured below, also see Table 5.5, Figures 5.20, 5.21 .

Team 1 FA 1: Lowest idle time, taskload count, teamwork count, and communication count.

Team 1 FA 2 This team has the highest mission, total task duration, and total busy time.

Team 2 FA 1: This team has the shortest mission, total task duration, and total busy time. Also has the highest teamwork count.

Team 2 FA 2: Has the highest idle time, taskload count, and communication.

Table 5.5: Total mission metrics for all teams.

Total Mission Metrics	High	Low
Total Mission Duration (s)	5841, T1-FA2	4109.1, T2-FA1
Total Task Duration (s)	12895, T1-FA2	8594, T2-FA1
Total Busy Time (s)	9745, T1-FA2	7695, T2-FA1
Total Idle Time (s)	10284, T2-FA2	6183, T1-FA1
Total Failure Time (s)	n/a	n/a
Total Taskload (count)	185, T2-FA2	162, T1-FA1
Total Teamwork Actions (count)	81, T2-FA1	58, T1-FA1
Total Physical Actions (count)	35, all	35, all

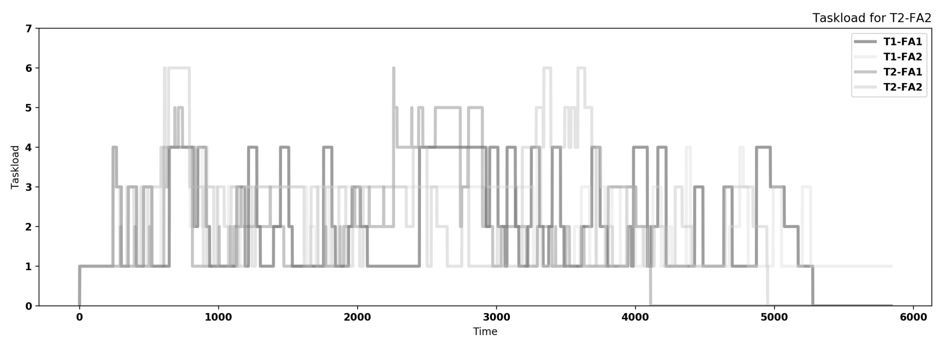


Figure 5.20: Taskload graph comparison for all teams.

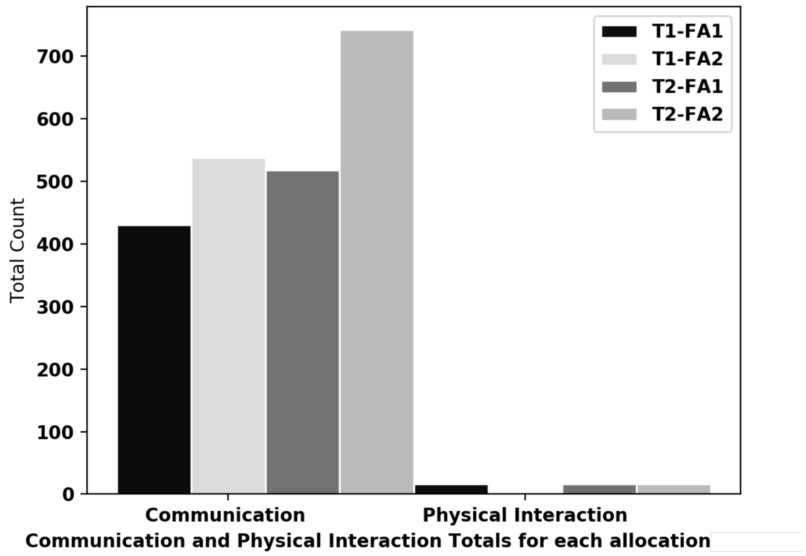


Figure 5.21: Communication and Physical Interaction cross team comparison.

Given these early observations, it may appear the Team 2 FA1 may be the most effective team. However, we can better our observations by comparing teams across main mission metrics, taskwork actions, and teamwork actions. Overall main mission metrics show that Team 2 FA 1 has the second highest idle time, but performs much better with the lowest busy and total mission duration time, Figure 5.22.

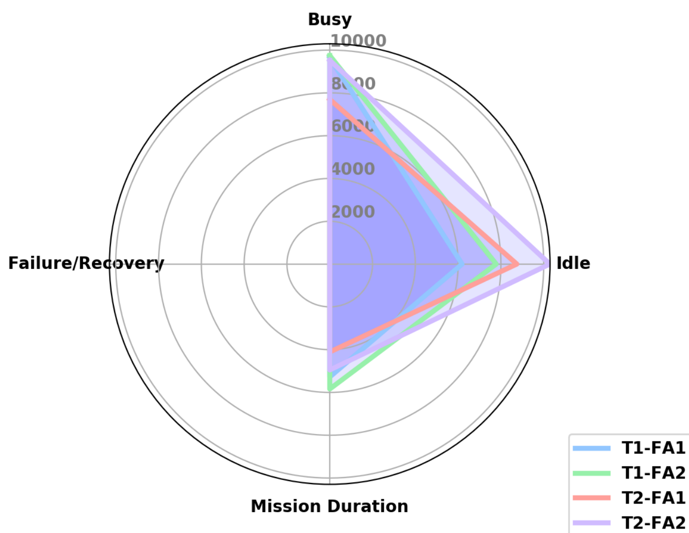


Figure 5.22: Main Mission Metrics: Time spend busy, idle, failure recovery, and total mission duration across all teams.

Alternatively, Team 1 FA1 has the lowest idle time but is among the highest for total busy time and total mission duration. The differences between the distribution of taskwork actions seem to be minimal, although we can see that Team 1 FA1 has the lowest count for most actions compared to Team 2 FA2 being the highest, Figure 5.23.

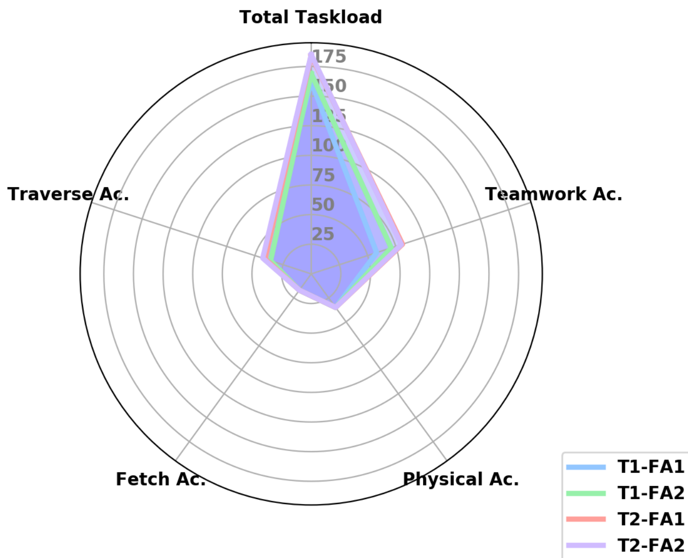


Figure 5.23: Taskwork Action Metrics: Total taskload count and spread across different actions (ac) for all teams.

Still the most interesting is the spread of teamwork actions as each team has unique differences. Team 1 FA1 and Team 1 FA 2 both require joint, control, and monitoring actions, while the former team requiring more joint and the latter more monitoring. Team 2 FA1 contains all teamwork actions but joint and Team 2 FA 2 contains only monitoring and control actions, Figure 5.24.

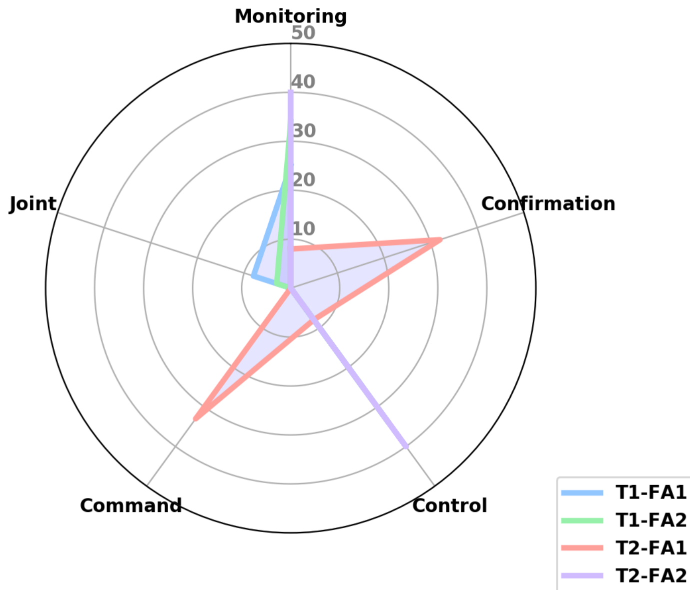
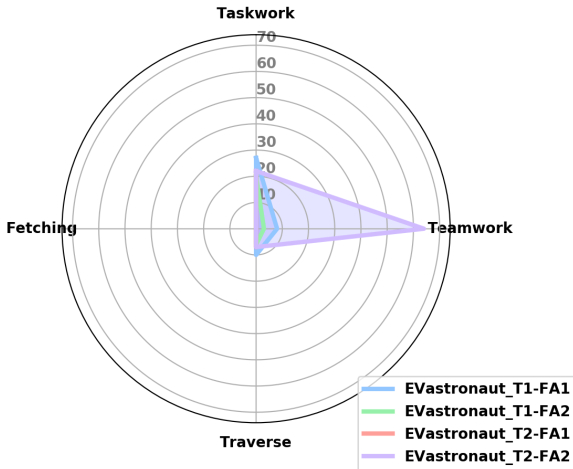


Figure 5.24: Teamwork Action Metrics: Division of work for teamwork actions for all teams.

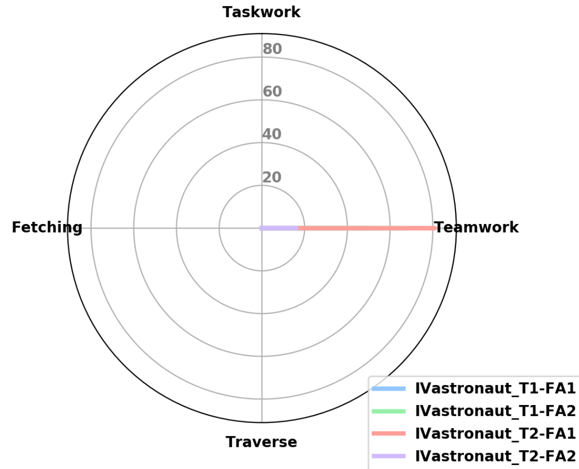
Earlier we mentioned that another perspective of evaluating simulated teams may be to look at the individual performance of members that play a key role within the team. In Figures 5.25 and 5.26 we can review the taskwork and teamwork performance of the two humans that work in all four team examples, IV and EV astronaut. We summarize the key differences in the following:

EV astronaut

- Taskwork: For EV, the most notable shift if the teamwork actions which spikes in Team 2 FA2. The other FAs contain similar levels of taskwork distribution, Figure 5.25a.
- Teamwork: In Team 2 FA 1, we find EV to not perform any teamwork actions. In both Team 1's FAs, we see the only teamwork action performed by EV are joint actions. However, Team 2 FA 2 sees an extreme spike in teamwork actions performed, being monitoring and control, Figure 5.26a.



(a) Taskwork Actions for EV Astronaut for all teams.

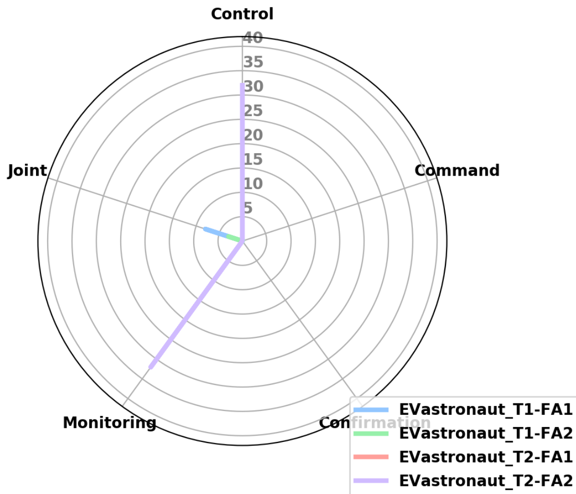


(b) Taskwork Action for IV Astronaut for all teams.

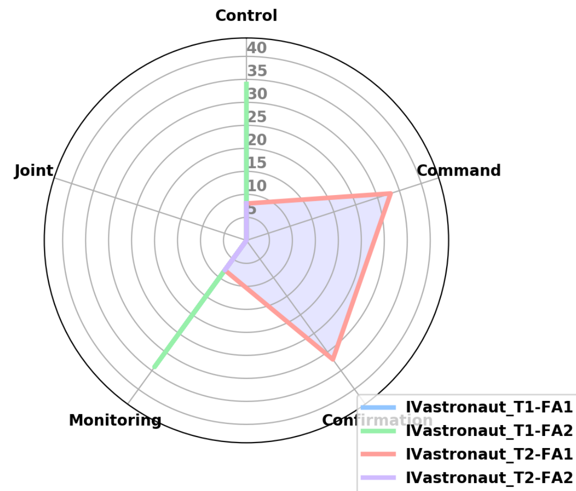
Figure 5.25: Taskwork Action breakdowns for EV and IV astronaut across all teams.

IV astronaut

- Taskwork: IV performs teamwork actions in both team FAs (see Appendix). Team 2 FA 1 has the highest Teamwork actions performed, while Team 2 FA2 has the lowest, Figure 5.25b.
- Teamwork: Team 2 FA 1 contains the most varied action distribution for IV (all but joint), while all other teams configurations include monitoring and control, Figure 5.26b. Interestingly Team 2 FA1 has a low amount of control and monitoring actions, while Team 1 FA 2 has the most of those two teamwork actions.



(a) Teamwork Action for EV Astronaut for all teams.



(b) Teamwork Action for EV Astronaut for all teams.

Figure 5.26: Teamwork Action breakdowns for EV and IV astronaut across all teams.

Each breakdown of teaming cross-teams can provide an interesting analysis of team design be it from overall team performance, per mission, or per agent perspective. For overall mission metrics, depending on the design concerns, i.e a higher or lower idle time, the preference for team configuration and task allocation may differ. When we look at mission specifics like task type (if taskload is a key concern) there may be again different preferences. Additionally, the load of each type of taskwork can be considered, i.e. fetching actions require more coordination than traversal tasks. The teamwork action distribution is yet another example of design preferences and evaluation of the true workload on each agent. Perhaps joint actions are preferred or to be avoided, or monitoring a robot is more distracting for individuals than directly controlling them is. If we dig even deeper into individuals themselves, there are other distinctions and considerations to be made, given their roles or capabilities.

5.4 Closing Thoughts

Through the expansion of work allocation metrics and modeling of human-robot team interaction, we are able to show how these various metrics allow us to get a more detailed and well-rounded understanding of human-robot teams. The modeling not only creates a more realistic simulation but also allows for more detailed metrics analysis. These expanded metrics provide a means to further study various areas of interest to the wider robotics and HRI communities through a work allocation perspective. In turn, we can answer questions that differ from traditional mission performance to address human-robot team concerns for improving individual performance, team member interaction, and more. Combined with high-level mission performance metrics, the interaction metrics allow for in-depth investigations of teaming.

The analysis for each team reveals a complex story spanning different levels of abstraction. The mission metrics provide a detailed picture of the performance and efficiency of the entire team for the defined mission. Mission level metrics are useful to determine if the current allocation of functionality and the team members and associated abilities are sufficient to meet the mission objectives. However, as our modeling revealed, there are more choices to be made in addition to the allocation of functionality. Specifically, the teamwork mode may also play an important role in the team's ability to meet mission objectives. Capturing the teamwork mode and analyzing the mission by cognitive step can highlight the differences between the team and work allocation options, predict situations of potential overload on individual agents, or locate information or resource choke points.

These team analyses also identify how a definition of a successful team can differ based on the goals of the mission and team designers. While one may not be unambiguously better or worse, these case studies reveal that assuming team effectiveness from regular mission metrics is not enough to judge a human-robot team's interactions and relationships. Instead, this work demonstrates how understanding both mission metrics, interaction, and individual performance can better describe human-robot team performance. Work allocation here is a key method for evaluating how team configuration and task assignment can impact the overall mission and teamwork metrics.

More specifically, the analysis also uncovers additional research questions with regard to thresholds and limits on individual agents – what is too much, what is too little? When is there too much reliance on a single agent? How should designers balance the cognitive vs. physical load on agents? Where are opportunities for robot agents to provide more or less support for humans? When to allow a performance or efficiency penalty to spread the work more evenly, or to provide a higher level of resilience? This work did not attempt to answer any of these questions directly, but by virtue of now computing these more detailed metrics, we can begin to ask and answer these questions.

Given that these metrics span a wide range of perspectives that designers may not fully take into consideration, more work should be done to observe the full impact that design decisions can have on teamwork and mission success. The proposed set of expanded metrics still leaves room for future improvement and additions. Additionally, simulations cannot cover all aspects of HRI like tiredness, frustration, boredom or others; therefore, testing these metrics for real human-robot teams working together in a human-in-the-loop study would provide realistic results for expected mission performance. Future work may involve more complex team examples, more detailed modeling of human-robot team interactions, and validation of these case studies to ensure that these metrics are valid representatives of real human-robot teams.

Although a breadth of interesting metrics can help inform design decisions, further analysis of different teams, member configurations, and teamwork modes is necessary to provide a more complete picture of how these design decisions impact metrics of human-robot teams. This work expands on the metrics that should be further explored for evaluating human-robot teams from a human factors perspective. Small changes in team assignment and structure are able to impact the team performance drastically and in ways that may be unexpected. To investigate human-robot teams more deeply there should be full consideration of mission performance, interaction, and impact on agents so that designers can thoroughly investigate and evaluate key components of human-robot interaction.

CHAPTER 6
THE SIGNIFICANCE OF MODELING TEAMWORK AND WORK AND VALIDATING SIMULATED HUMAN-ROBOT TEAMS

As human-robot teams become more prevalent in complex and safety-critical work environments, so does the need for methods to effectively design them [50]. Important aspects of their design include effectively allocating the required work activities to different agents, humans and robots, within the team; and establishing effective teamwork protocols for human-robot interaction. These components are also key considerations within our modeling and design of teams in our simulation. To verify our simulation, we asked if our model fo human-robot teaming in nominal and off-nominal scenarios demonstrate an appropriate impact of teamwork metrics. We address this through investigating the effect of failures and how they propagate throughout the simulation through agent action traces and quantitative metrics.

This section describes a HITL experiment that refines predictions of human-robot teaming provided by computational simulation. Computational modeling can describe (1) the dynamics of the work a team is expected to collectively perform; (2) the team performance impact of allocating different work activities to different agents, human and robot; and (3) the team performance impact of different 'teamwork' protocols for human-robot interaction. Computational predictions of the relative effects different work allocations and human-robot teamwork protocols were confirmed by the experiment. Further, bootstrapping analysis demonstrated how the computational models can be further improved from HITL results, both in terms of refining estimates of specific activities and in terms of identifying other important effects to incorporate, such as communication times and the time to transition between actions.

6.1 Verifying Teaming Constructs in WMC

Much of this work so far has explored modeling various types of human and robot teams in deep space scenarios. In particular, we were interested in studying human-robot teams operating in off-nominal situations. Due to the ability of modeling and simulation as a useful tool to evaluate systems early-in design, we modeled the descriptions of robotic failures in human-robot teams.

This case study assesses the capabilities of four different robots and two humans in an in-orbit maintenance mission and focused on exploring failures and responses to interruptions in teaming [118]. In this work, we explore situations in which robots fail to correctly complete some of their assigned actions. These failures had two variations, 1) the action that failed being interdependent vs. independent and 2) robot awareness of failure, whether or not they were aware of failing the action or not. The goal of this work was to verify that modeling and simulation is capable of effectively evaluating responses to failures.

This following verification work describes a methodology for simulating failures and investigating the impact on metrics that measure human-robot teams. We investigated teams of astronauts and a diverse range of space robots working together to jointly complete a standard inspection and maintenance EVA operation. In particular, we identified key factors in human-robot work allocation that impact communication, fault identification, and taskload on human and robot agents alike. These result in fundamental tradeoffs in nominal and off-nominal scenarios for teamwork and mission metrics like mission duration, task duration, workload for human and robot agents, and communication time and efficiency.

6.1.1 Metrics of Off-Nominal Failure Scenarios

WMC is capable of capturing a wide range of metrics that define taskwork efficiency and performance as well as teamwork and communication:

- Taskload and timing of executing tasks can be evaluated for the taskload of agents in the aggregate and at any point in time.
- Performance measures include mission duration, total idling time, and total time on tasks, and allow for quantitative analysis and comparison between varying failures. These measures in this case study are normalized against a baseline scenario with no failures.
- Information transfer between agents is logged as communication when agents get a resource that was set previously by another. The information transfer can be logged overall, recording the complete set of information requirements. Furthermore, it can be categorized according to the communication channel by which the agents would communicate this information.
- Average action timing compares how long failed actions take to complete compared to the nominal case.

6.1.2 Scenario Cases Studies for Failures

This scenario has a human-robot team conducting routine maintenance on the exterior of the spacecraft. Specifically, the team inspects ten items (three space panels, four solar panels, and three filters) and replaces them if needed. The agents working in this scenario include two human astronauts (extra-vehicular [EV] and intra-vehicular [IV]), a remote manipulator system (RMS), two humanoid robots (Humanoid Robot), a fetching robot, and Mission Control Center (MCC). The robotic agents can be controlled directly or through commands by human operators; this assignment depends on their capabilities for the action in question and allocation of the task.This scenario requires the agents to traverse to locations specific to each action, the panel and filter layout are in Figure 6.1.

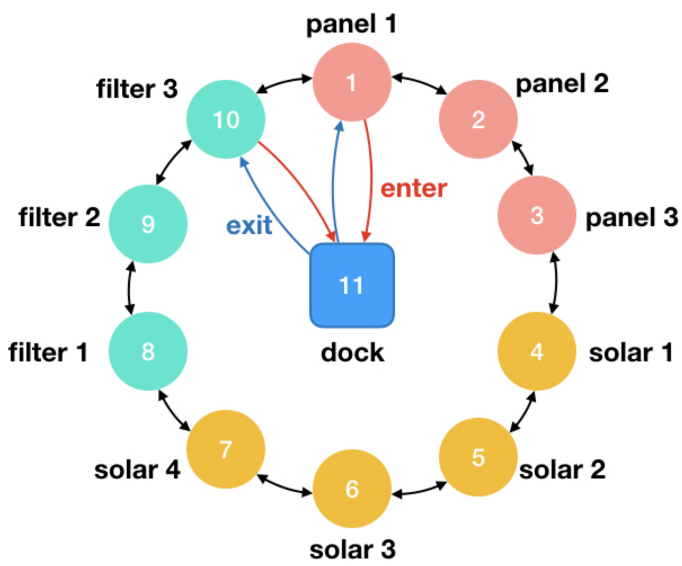


Figure 6.1: Map of location on exterior of spacecraft.

IV has the majority of channels, with access to EV, Humanoid Robot, RMS, and the fetching robot. EV mainly communicates with the Humanoid Robot and the Fetching robot. The Humanoid Robot can also talk to the Fetching robot. MCC is capable of speaking only to IV and EV astronauts through Radio1. This communication structure constrains which agents can respond to failures and the mechanisms by which they can confer on and interrupt each other Figure 6.2 .

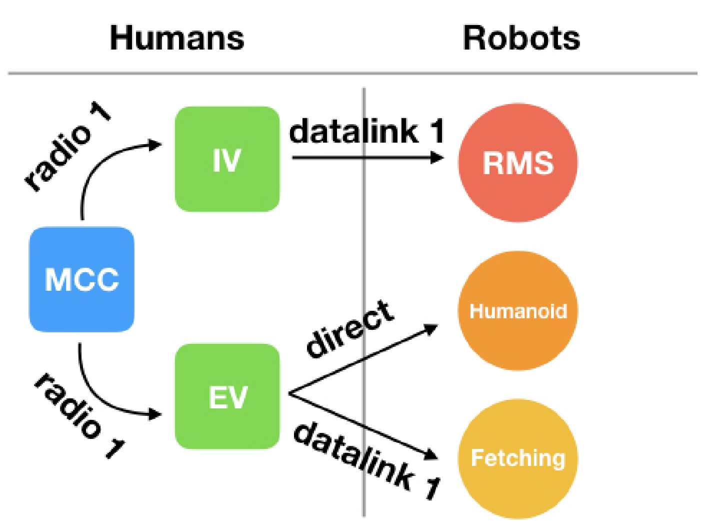


Figure 6.2: Communication channels for agents in simulation.

There are a few levels to our test cases. The first being which agent fails, as only one robot agent failed per case being either RMS or the Humanoid robot. After determining the agent to fail, we set the failure amount as no failure (nominal), one failure, or two failures. After determining the number of failures, we determine which task(s) will fail. We defined a set of actions that would fail: an inter-dependent action with other actions performed by other agents, action, and a simpler, independent action that is not time sensitive and that no other actions are dependent on, see Figure 6.3 . Finally, we determine the awareness of failure (known or unknown). We ran the 14 different failure cases noted in Figure 6.4 to examine cases with either one or two failures.

HTA ID	Name	FA-RMS				FA-Robonaut			
		Authority/ Responsibility	Verfication	Control Mode	Failure	Authority/ Responsibility	Verfication	Control Mode	Failure
1.1.	Prepare	EV/EV	M	CS		EV/EV	M	DT	
1.2.	Leave dock	EV/EV	M	CS		EV/EV	M	DT	
2	Traverse	EV/EV, Robot*/IV	M	CS, DT		EV/EV, Robot*/IV	M	DT	
3.1.	Get inspection tools	EV/EV	M	CS		**Humanoid/IV**	C	DT	
3.2.	Apply inspection tools	EV/EV	M	CS		**Humanoid/IV**	C	DT	**FAIL**
3.3.	Store inspection tools	EV/EV	M	CS		**Humanoid/IV**	C	DT	
4.1.	Get repair tools	EV/EV	M	CS		**Humanoid/IV**	C	DT	
4.2.a.	Get new panel	**RMS/IV**	C	CS	**FAIL**	RMS/IV	M	DT	
4.3.a.	Remove broken panel	EV/EV	M	CS		EV/EV	M	DT	
4.4.a.	Dispose of broken panel	**RMS/IV**	C	CS		RMS/IV	M	DT	
4.5.a.	Emplace new panel	EV/EV	M	CS		EV/EV	M	DT	
4.2.b.	Get new filter	**RMS/IV**	C	CS		RMS/IV	M	DT	
4.3.b.	Remove old filter	EV/EV	-	CS		EV/EV	M	DT	
4.4.b.	Dispose old filter	**RMS/IV**	C	CS	**FAIL**	RMS/IV	M	DT	
4.5.b.	Emplace new filter	EV/EV	-	CS		EV/EV	M	DT	
4.2.c.	Replace wiring	EV/EV	M	CS		EV/EV	M	DT	
4.6.	Store repair tools	EV/EV	M	CS		**Humanoid/IV**	C	DT	**FAIL**
5	Enter dock	EV/EV	M	CS		EV/EV	M	CS	

Figure 6.3: Failure Test Cases. a. Function allocations are specified as ¡authorized agent¿/¡responsible agent¿. b. Verification can be: M = monitoring, C = confirmation. c. Control mode can be: NC = no control, DT = direct teleoperation, CS = command sequencing.

- RMS Failures
  - Get New Panel (inter-dependent): a repair which completes the replacement of a broken panel by a new panel) robot fails to get a new panel. As a result, the repairs for this broken panel cannot continue.

- Dispose Old Filter (independent): disposing old filter to trash robot fails to dispose of the filter.

- Humanoid Robot Failures
  - Apply Inspection Tools (inter-dependent): Apply the tools to perform the panel inspection and provide the damaged panel information robot fails to apply tools properly. Similarly, other panel repair actions are postponed until this action is completed.
  - Store Repair Tools (independent): Put away the repair tools robot fails to store tools.

		Function Allocation		
FA RMS (EV, IV, RMS, Fetching, Robonaut)	No failure	1 failure		2 failure
		Known: get	Known: Dispose	Known: get, Unknown: dispose
		Unknown: get	Unknown: Dispose	Unknown: get, Known: dispose
FA ROB (EV, IV, RMS, Fetching, Robonaut)	No failure	1 failure		2 failure
		Known: apply	Known: Store	Known: apply, Unknown: store
		Unknown: apply	Unknown: store	Unknown: apply, Known: store

Figure 6.4: 14 test cases

6.1.3 Case Study Discussion

Our results investigate the effect of failures on overall metrics (overall action duration total mission duration, total time spent on actions, total mission idle time) and teamwork metrics (total number of communications and communication channel usage) for both failure test cases. Additionally, we compare the agent action traces and metrics of each agent for the failure test cases and to the case of no failures.

- Robot Capabilities and Taskwork: Overall, it appears that the capabilities of the robot affect the communication or information. RMS robots are rooted in physical space and are assigned fewer traversals. Therefore, certain actions like traversals and fetching (and their addition to taskwork metrics) and communication between agents are directly impacted. Additionally, the robot's proximity to the action area can be more readily assigned to robots capable of locomotion. Depending on the scenario structure and taskwork, certain types of robot agents may get more work or must communicate more actions to human teammates. Being cognizant of a failure as a robot has a direct impact on communication load for robots who cannot move, and therefore cannot use their time to do other tasks that do not involve locomotion.

- Failure Timing: Failures can cause some reductions in metrics for total time, taskload, and idle time for the agents who re-do previously failed actions. This can occur with lower impact failed actions that occur in the beginning when there is less work, versus higher priority failed tasks that occur when many other agents are already busy doing other work. Failures that occur closer to the end of a task sequence can be more disruptive depending on the business of the agents in the simulation during that time , see Figure 6.5. When agents are potentially more spread out in various locations and engaged with their own other tasks, failures have a bigger impact on agents. This timing inadvertently affects agents who are not directly related to the failure, who, like the fetching robot, then get assigned new tasks for fetching and traversal to assist other agents. Failures can also cause increases in metrics for agents who re-complete failed actions.

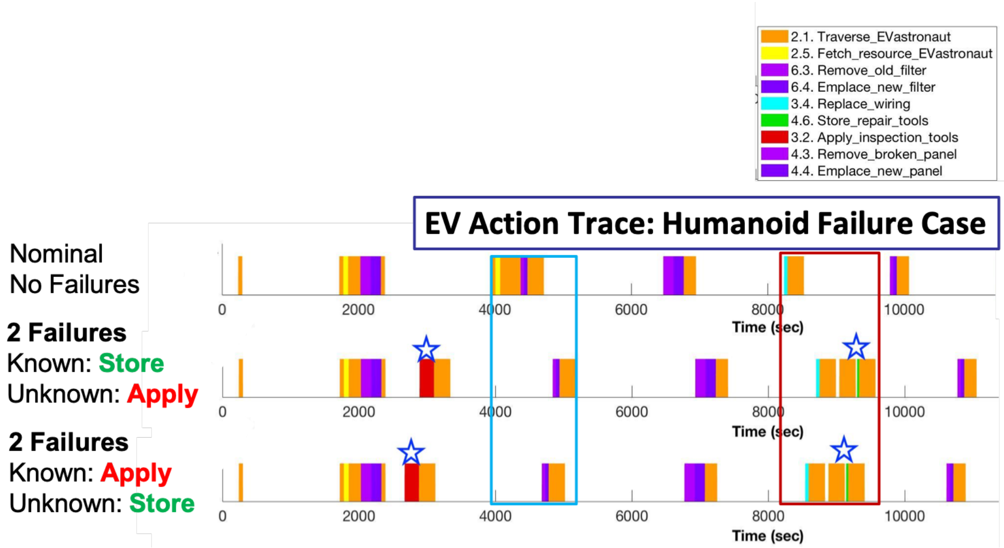


Figure 6.5: The impact that robotic capabilities of locomotion and failure timing have on taskwork timing for off-nominal comparisons to nominal cases.

- Location: when choosing between two agents in the simulation, the closer agent to the location of the action will get assigned the fetching and traversal actions. Given this, to lower taskload, teams should try to ensure traversal and fetching depending on the location of agents in the team. Location, and the capability of traversal, is also enmeshed with failure timing as the relative position of agents to failures can cause increases in taskload, mission time, and result in inefficiencies, see Figure 6.6.

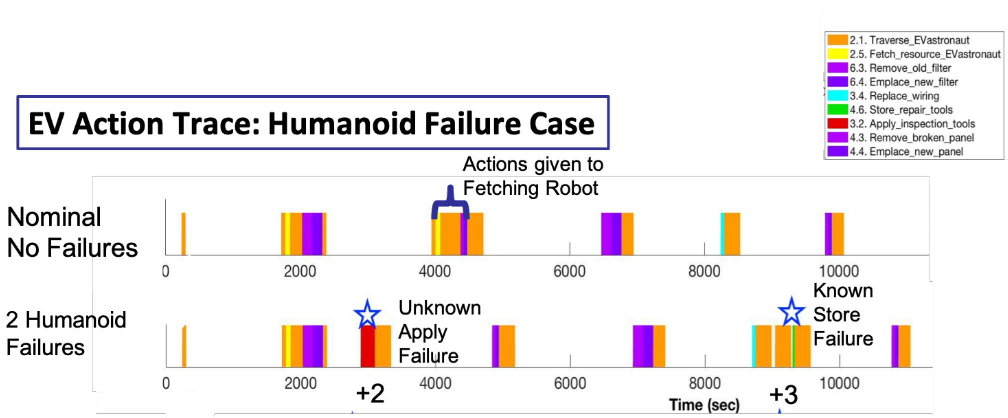


Figure 6.6: The impact that robotic capabilities of locomotion and failure timing have on proximity to taskwork in off-nominal comparisons to nominal cases.

- Idle Time: Can be affected from two perspectives, either less taskwork is assigned from the first and last action the agent completes (which affects the idle time since the last action may be a failed action), or by a factor of having more work assigned such that the agent is truly busier throughout the mission.

- Failure Knowledge: These robot's knowledge of their failures has less impact on the results of the scenarios than anticipated when looking at the idle time, taskload, and total time. However, when observing actions traces, we can see that unknown actions push back timelines more than known actions, due to the timing of failures. Failure actions also cause more taskload on agents that fail since they complete the actions incorrectly, as opposed to realizing the failure early-on. We can see that knowledge of failures has a positive impact on closeness to baseline values over unknown failures that may occur throughout the system for single failures that can occur.

- Action Dependency: High impact actions are inter-dependent actions, while lower impact actions were the independent actions. In this case study, we see that the action timing depends on action dependency, where inter-dependent actions tend to extend timelines more, lengthening mission time and agent idle time. This ixs because more actions must wait upon the failure being resolved, causing some agent to idle in the meantime. Action dependency is also dependent on failure knowledge, see Figure 6.7. Independent action failures are likely to have higher impact as future actions depend on completion of these actions.

Metrics	No Fail	1 Failed Action			
		Known		Unknown	
		Get	Dispose	Get	Dispose
Mission Duration (s)	6990	160	180	40	0
Total Idle Time (s)	19770	790	-60	710	90
Total Time on Task (s)	11550	0	0	240	220
Total Info Transfer (#)	148	8	3	11	8

Figure 6.7: The impact that action dependency and failure knowledge have on failure impact in off-nominal comparisons to nominal cases. Values shown as a difference between no failure case.

Through these observations, computational simulations of work allocation are able provide insight into human-robot teaming. This work investigated the impact of robotic failures (known and unknown) in human-robot teams. Their impact was examined in cases where the robot was aware of its failures and where the robot did not know, and therefore, did not respond to it. We also identified key factors that impact teamwork metrics.

Here we show that teamwork and interactions between team members, particularly communicating failures, can cause distinct changes in our system?s metrics. Failures can be task specific or be affected by a multitude of factors which impact metrics for taskload, mission time, total time spent on tasks, and information transfer. Failure timing (when agents are busy or idle), agent proximity to failures and future actions, failure knowledge, and inter-dependent actions (have following dependent actions) are key determinants of how the failure is resolved and how agent metrics are impacted. This experimentation helped verify the ability of our model to accurately simulate different human-robot teams in off- nominal scenarios, showing that we can build future models on top of this architecture that provides expected results from simulations. We were also able to show how small changes in the interaction can propagate differently depending on the context of the scenario. To further investigate human-robot teaming, later sections will describe refining and validating more metrics that measure a range of different aspects of teamwork.

6.2 Validating Teaming Metrics in WMC

Computational modeling and simulation can predict many measures of team interaction and performance, with the particular ability to capture emergent effects and complex interdependencies within the team's activities. However, while recent studies have demonstrated the ability of this approach to quickly compare different team designs, the computational models depend upon estimates of the duration of each agent's actions.

Thus, this section describes a Human-In-The-Loop (HITL) experiment designed to refine computational models supporting the design of human-robot teams. The experiment results are first examined directly for the extent that they confirm common hypotheses about human-robot team design. Then, an iterative bootstrapping technique elaborates on the experiment results to refine the computational model's predictions and help validate of specific aspects of our simulation metrics.

6.2.1 How to Validate WMC Models and Output

WMC is relatively unique in that it models the work separately from the agents. Thus, in simulating a team's work models the taskwork can be fluidly assigned to appropriate agents during run-time. Further, the WMC simulation framework can also identify during run-time when teamwork actions are needed. This structure allows for rapid evaluation of a wide range of different allocations of the taskwork between the team's agents, and of different teamwork protocols such as different modes for human-robot interaction. These two aspects of WMC are key to both modeling considerations and metric evaluation for human-robot teams.

In this work, WMC has been applied to study work allocation in human-robot teams for manned spaceflight operations, the inspiration for the experiment described in Section 5 [118, 51, 52]. Work allocation is the set of design decisions defining how work is distributed among team members. [33]. Past research in work allocation has focused on various perspectives of agent and team performance. Such metrics include total task duration, human intervention time, wait or idle time, resource consumption, or task-specific performance measurements and workload estimates [38, 37, 39].

The raw results of a WMC simulation comprise a detailed timeline of all the taskwork and teamwork actions completed by each agent, and the information and resources that each action 'gets' and 'sets.' From these raw results, many more-aggregate measures can then be calculated. Some metrics span the entire team and may reflect measures of mission performance, such as total mission duration. Others provide more detailed assessments of the time that any (or all agents combined) are busy or waiting on others. These results can be further extrapolated to form estimates of aggregate taskload, duration of periods of taskload saturation, or use of consumables such as, in space missions, spacesuit oxygen. Where the types of activities are important, the relative time and amount of taskwork versus teamwork can be compared, or specific aspects of human-robot interaction identified such as the number of times a human needs to monitor the robot. Each agent's requirements for information and resources can be recorded, as can the need for inter-agent communication to meet these requirements.

However, these computational simulations are sensitive to estimates of some parameters describing the work. Before a team is formed, work models can use best-estimate assumptions of the duration of specific actions, for both taskwork actions and for teamwork actions. Simulations of these initial models will have a good first-principles assessment of the team's action *sequence*, but corresponding estimates of mission duration will be sensitive to the accuracy of the assumed action durations within the mission. Thus, this work examines how the models can then both (1) serve as input to HITL experiment designs by identifying key aspects of the human-robot team design meriting further evaluation, (2) be refined by the HITL experiment results, particularly in better parameterizing estimates of time and duration within the model, and (3) validate the keystones of WMC modeling and metrics by testing teamwork modes and taskwork allocation.

6.2.2 HITL Experiment Design

The experiment task was crafted to represent an extra-vehicular mission in outer-space orbit where a team of three (an intra-vehicular [IV] astronaut, extra-vehicular [EV] astronaut, and a Turtlebot) worked together to maintain the spacecraft's exterior. This mission has also been modeled in WMC in Chapter 5.

Experiment Task

Two different locations needed to be inspected, one with a panel and the other with wiring component. If inspection found deterioration in either component, it needed to be replaced (panel, using tools) or repaired (wiring). The tools were located at a third location (Dock), where three locations were about 5 meters apart from each other requiring movement between each. Figure 6.8 shows the wiring unit, panel, and tools.

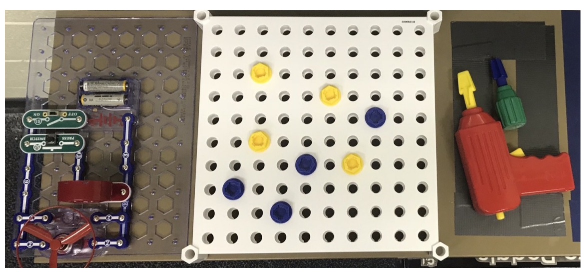


Figure 6.8: Experiment Setup

The IV (experiment supervisor) coordinated the activities for each run using a script. The robotic agent was represented by a Turtlebot. This Turtlebot was able to traverse between locations to fetch tools and, for other tasks that it could not complete, an experimenter Wizard-of-Oz'd the Turtlebot's abilities.

Participants assumed the role of the EV astronaut and worked together with the Turtlebot robot in one room, communicating with the IV astronaut in another room, to complete all the maintenance tasks. Each run lasted about 30 minutes.

Independent Variables

Reflecting important questions in work allocation and human-robot interaction in space-flight missions, the independent variables for the main body of the experiment were:

- Task allocation, denoting the functions performed by the robot:
  1. Inspect (I): The Turtlebot appeared to the participant to inspect the wiring, with the results of the inspection communicated to the participant through the same iPad used to communicate with the IV. If the inspection resulted in a broken panel, the participants would have to fix the panel.
  2. Fetch and Inspect (FI): With this task allocation the Turtlebot additionally supported the participant by fetching and delivering tools and panels.
- Teamwork Mode:
  1. Control/Confirmation (CC): The participant overheard, and thus was aware of, the IV's actions in controlling and confirming every action the Turtlebot conducted.
  2. Monitoring (M): The implied teamwork mode was that the IV was silently monitoring the robot as it acted independently, without the participant overhearing on-going command and confirmation.

Thus, combining the task allocations with teamwork modes in a 2x2 design created four nominal conditions: FI-CC, FI-M, I-CC, I-M. Every participant completed each of the four nominal cases, where the order of runs was designed using a Latin square to prevent carry-over effects.

A fifth run after the nominal conditions represented an off-nominal case where the robot reported, mid-task, that it could not inspect the panel; the participant then had to inspect the panel directly.

Table 6.1: Independent Variables

Condition	Description	Hypothesis
FI	Turtlebot can fetch tools and inspect panels	Robot may be perceived as more helpful
I	Turtlebot can inspect panels only	Robot be perceived as more idle and less helpful
CC	IV must command and confirm robot's actions	Team may seem more communicative/provide more situation awareness
M	IV must monitor robot's actions	Robot may seem more capable/efficient. EV may not be as aware of robot
Failure	Same as nominal but robot fails one inspection action	May exaggerate the findings in the nominal cases

Table 6.2: Scenarios

Function	FI-CC			FI-M			F-CC			F-M		
	EV	IV	T	EV	IV	T	EV	IV	T	EV	IV	T
Inspection	✓		✓	✓		✓	✓			✓		
Panel Rpl.	✓		✓	✓		✓	✓		✓	✓		✓
Wiring Rep.	✓			✓			✓			✓		
Fetching	✓		✓	✓		✓	✓			✓		
Ctr/Conf		CC						CC				
Monitoring					M						M	

Note: EV = Extra-vehicular astronaut (Participant), IV = Intra-vehicular astronaut; T = Turtlebot Robot, M monitoring, CC = command sequencing with confirmation.

Procedure

The experiment consisted of an introductory briefing, collection of informed consent, a training mission, a data collection set, (with questionnaires after each of the five runs), and a post-experiment questionnaire. The introduction and briefing explained the mission and overall process. The training mission was consistent across participants, familiarizing participants with the tasks to be performed. The measurement phase consisted of the five measurement runs four nominal cases and one failure case); breaks were given after each run if needed. After each run, participants were asked to complete a questionnaire. Following all runs, they were asked to complete a final questionnaire regarding their workload, robot interactions, and teamwork.

Participants

Twenty-four college students from the Georgia Institute of Technology participated. Participants were not required to have any experience with robots but were required to have proficiency in speaking and reading English.

Dependent Measures

During each run, we tracked the timing (start and stop) of both the participant's and robot's actions to capture their duration, reported into the same format as WMC raw results. Using this information, we tested for specific HRI metrics of interest for mission outcome and workload including total mission duration and its constituent parts: busy time, idle time, and traversal time. These total times were captured for each team members and also summed across all team members.

Additionally, questionnaires after each run sought participants' perception of their workload using the 6 NASA TLX rating scales [146], and their perception of the robot. At the end of the experiment, a final questionnaire asked participants to rate their experience with the robot and to rank each experimental condition (FI-CC, FI-M, I-CC, and I-M) as to which provided the most effective robot teammate and their perceived idle time.

6.3 Experiment Results and Comparison to Computational Simulation Predictions

As described in the following sub-sections, two analyses were conducted. First, analysis of the participants' questionnaire responses characterized how participants perceived the different human-robot conditions created within the experiment (Section 6.3.1). Then, an iterative bootstrapping method is applied to refine the assumed action durations and validate the resulting teaming metrics within the computational model. (Section 6.3.2).

6.3.1 Participant Perceptions of the Human-Robot Team

After each run, participants indicated their immediate perception of that run's task allocation and teamwork mode on a questionnaire. One aspect of this questionnaire asked participants to rate their workload on the six sub-scales used in the NASA TLX. For these six measures a two-way, within-subjects ANOVA examined the effects of the independent variables (task allocation (FI,I) and teamwork mode (CC,M)), with commensurate assumptions tested by Shapiro-Wilk's normality checks and Levene's homoscedasticity checks.

The results are summarized in fig. 6.9 and in Table table 6.3. Mental demand, physical demand, and effort were found to have statistically significant results between the task allocations. When the robot was fetching and inspecting (FI) versus only inspecting (I), participants reported lower levels of mental and physical demand and effort . No significant interaction effects were found between the task allocation (FI,I) vs teamwork mode (CC,M) in any measure.

Table 6.3: Work allocation (FI vs I) Two-way Repeated Measures ANOVA table

NASA Scale	f	η^2	$\mu(I - FI)$	p
Mental*	7.024	0.251	0.574	0.015
Physical*	8.574	0.290	0.696	0.008
Temporal	0.700	0.036		0.413
Effort*	16.454	0.428	0.893	0.001
Frustration	0.176	0.009		0.680
Performance	3.019	0.121		0.096

*indicates $p \leq 0.05$

means scaled from 0 to 1

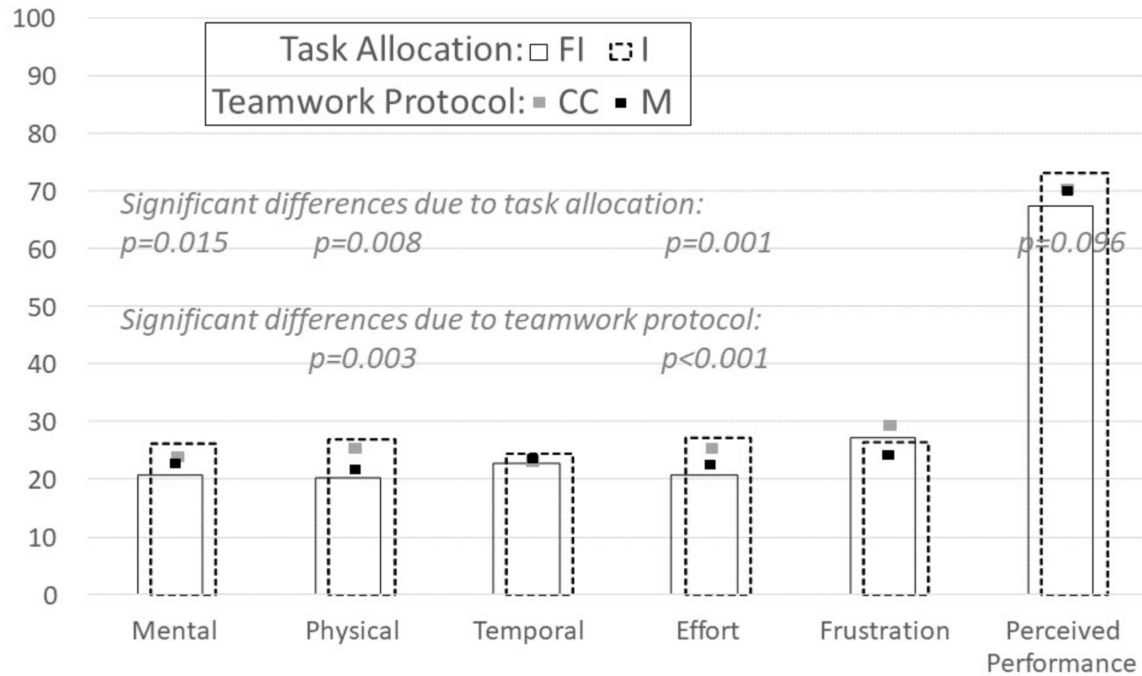


Figure 6.9: Average participant ratings in the six NASA TLX sub-scales with the different task allocations (FI,I) and teamwork protocols (CC,M)

For teamwork modes, the participants rated the command and control (CC) teamwork mode as having significantly more physical work and taking more effort than the monitoring (M) mode, see Table table 6.4. This is surprising given that the only difference in the participants' experience between these two modes was whether they overheard the IV controlling the robot, without needing to perform any extra duties themselves.

Table 6.4: Teamwork Mode (CC vs M) Two-way Repeated Measures ANOVA table

NASA Scale	f	η^2	$\mu(CC - M)$	p
Mental	0.880	0.027		0.456
Physical*	11.051	0.345	0.547	0.003
Temporal	0.000	0.000		0.994
Effort*	20.774	0.486	0.725	<0.001
Frustration	1.690	0.078		0.208
Performance	0.266	0.012		0.611

*indicates p≤0.05

means scaled from 0 to 1

When asked to rate their interactions with the robot after each run, no differences were found between participants' ratings as a factor of teamwork mode (CC versus M) (fig. 6.10, table 6.6). This is reasonable considering these teamwork modes did not directly impact the EV astronaut, as the IV was responsible for commanding, monitoring, and confirming the Turtlebot. Additionally, the taskload and resource exchanges were the same within the teamwork mode condition. Idle time, however, was higher for command and control compared to monitoring, particularly for the FI work allocation.

On the other hand, participants rated the Fetch and Inspect (FI) allocations consistently better than Inspect (I) allocations (fig. 6.10, table 6.5). To put these ratings in context: In FI conditions, the average participant wait time was 312 seconds (while the robot fetched tools), but their idle time with the I conditions was only 161 seconds as the participants instead needed to fetch their own tools. This means that in FI cases, where the participant waited for almost double the amount of time in I cases, they perceived the robot as a more effective teammate (Question 2). Therefore, the fetching ability of the robot was a factor in participants rating it as a more effective teammate even though its slow traversal speed increased the amount of wait time for the participant. In addition, the number of resource exchanges for FI allocation was 4 for fetch and inspect and none for only inspect. This matches the participants responses to Question 1 where they perceived more interactions in the inspect case.

In FI allocations, the total participant taskload was 24 while it was 33 in I cases. This matches the participants responses to Question 3 where they felt the robot in FI runs significantly reduced the taskload compared to I runs. Additionally, giving the robot the capability to fetch and inspect increased the perception of effectiveness but came at the cost of increased attention to the robot (Question 4).

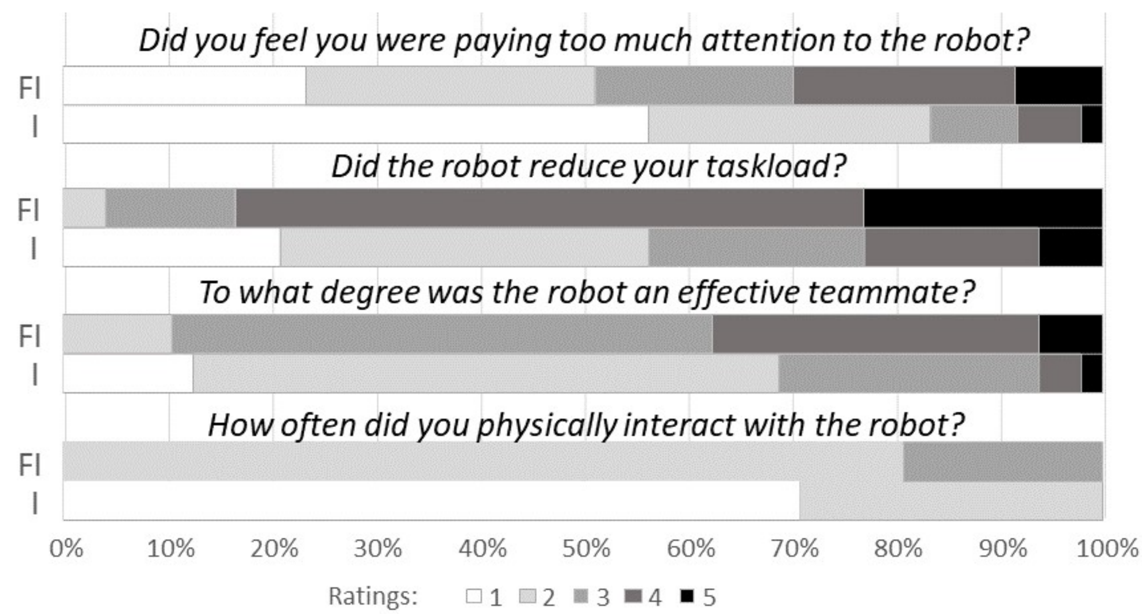


Figure 6.10: Percentage of participants' ratings on 4 questions of robot effectiveness, separated between FI and I task allocations

Table 6.5: Work Allocation Comparison for Questionnaire

Question	Allocation†	p
1. How often did you physically interact with the robot?*	FI (66.81)	
	I (30.19)	<0.001
2. To what degree was the robot an effective teammate in this mission?*	FI (64.19)	
	I (32.81)	<0.001
3. Did you feel the robot helped to reduce your taskload?*	FI (64.70)	
	I (32.30)	<0.001
4. Did you feel you were paying too much attention to what the robot was doing?*	FI (58.85)	
	I (38.15)	<0.001

*indicates p≤0.05, † is defined by the mean rank, rank scaled from 0 to 100

Table 6.6: Teamwork Mode Comparison for Questionnaire

Question	Teamwork†	p
1. How often did you physically interact with the robot?	CC (48.18) M (48.82)	0.898
2. To what degree was the robot an effective teammate in this mission?	CC (48.78) M (48.22)	0.917
3. Did you feel the robot helped to reduce your taskload?	CC (48.08) M (48.92)	0.879
4. Did you feel you were paying too much attention to what the robot was doing?	CC (48.63) M (48.38)	0.963

*indicates p≤0.05, † is defined by the mean rank, rank is scaled from 0 to 100

Finally, at the end of the experiment, the participants were asked to rank from best to worst the four different team conditions in terms of which provided a robot serving as the most effective team mate. A Wilcoxon rank sum test was used to analyze the rankings participants gave each work allocation. We found the FI-M and FI-CC conditions, in which the robot could fetch, were rated as significantly better than the lowest rated condition (I-CC). The same statistical effects were found in participants' rankings of the conditions in terms of the amount of time the participant was idle during a mission. However, there was no significance found in rankings of the conditions based on the participants' assessment of the time efficiency afforded by the allocation.

Table 6.7: Comparison of All Missions

Question	FA	Mean Rank	Significant Differences
Rank the missions where the robot was the most effective teammate	FI-CC	2.17	FI-M - I-CC (0.037)
	FI-M	2.17	FI-CC - I-CC - (0.026)
	I-M	2.71	
	I-CC	2.96	
Rank the missions according to how much idle time you experienced	FI-M	2.21	FI-M - I-CC (0.039)
	FI-CC	2.29	FI-CC - I-CC - (0.040)
	I-M	2.54	
	I-CC	2.96	
Rank the mission according to how time efficiently the team was working	FI-CC	2.29	no significance found
	FI-M	2.42	
	I-CC	2.58	
	I-M	2.71	

lower rank is better, ranked from 1-5

6.3.2 Refining and Validating the Computational Model with HITL Experiment Results

To improve the computational model's assumed action durations using the experimental data, without risking overfitting the data, we chose a bootstrapping approach. It is a resampling technique where a subset of data is chosen to calculate a given statistic. We used this technique to compare total mission duration, total idle time, participant busy time, participant idle time, and participant traversal time. WMC takes in task durations as an input, thus bootstrapping resampled a subset of all participant data to reflect each action duration. Bootstrapping then compared the average values to provide accurate timings for WMC.

Specifically, here we will define all participant data as $p \in P$ and n total bootstrapping iterations. The two output sets are defined as WMC data $w_i \in W$ and averaged validation participant data $v_i \in V$. Each bootstrapping iteration involved the following three steps:

1. Sample percentage of participant runs P into subset S so that $S \subset P$
2. Average action durations across a subset S and run WMC with durations saved into output w_i

3. Average set $P - S$ participant data metrics into singular v_i

This procedure was repeated for n iterations to produce set W consisting of WMC output and set V consisting of averaged participant data collected in the HITL experiment. The size of the output sets therefore were of length n ($|W| = |V| = n$). Here, we chose $n = 1000$ iterations and sampled 80% of P every iteration, using these results to refine the WMC model parameters. Participant runs with action durations above or below 1.5 times the interquartile range were considered outliers and removed from the sampling process. The remaining 20% of P was then used to assess the accuracy of the parameterized WMC model results.

Mission Metrics Across the Team

These mission metrics across the team are total mission duration, total busy time, and total busy time. For total mission duration, after the bootstrapping was used to parameterize the action durations, WMC accurately predicted the same trend in mission duration found in the HITL experiment (fig. 6.11): the team conditions are correctly predicted by WMC from fastest to slowest being I-M < FI-M < I-CC < FI-CC. The total mission duration, however, is consistently under-estimated by WMC.

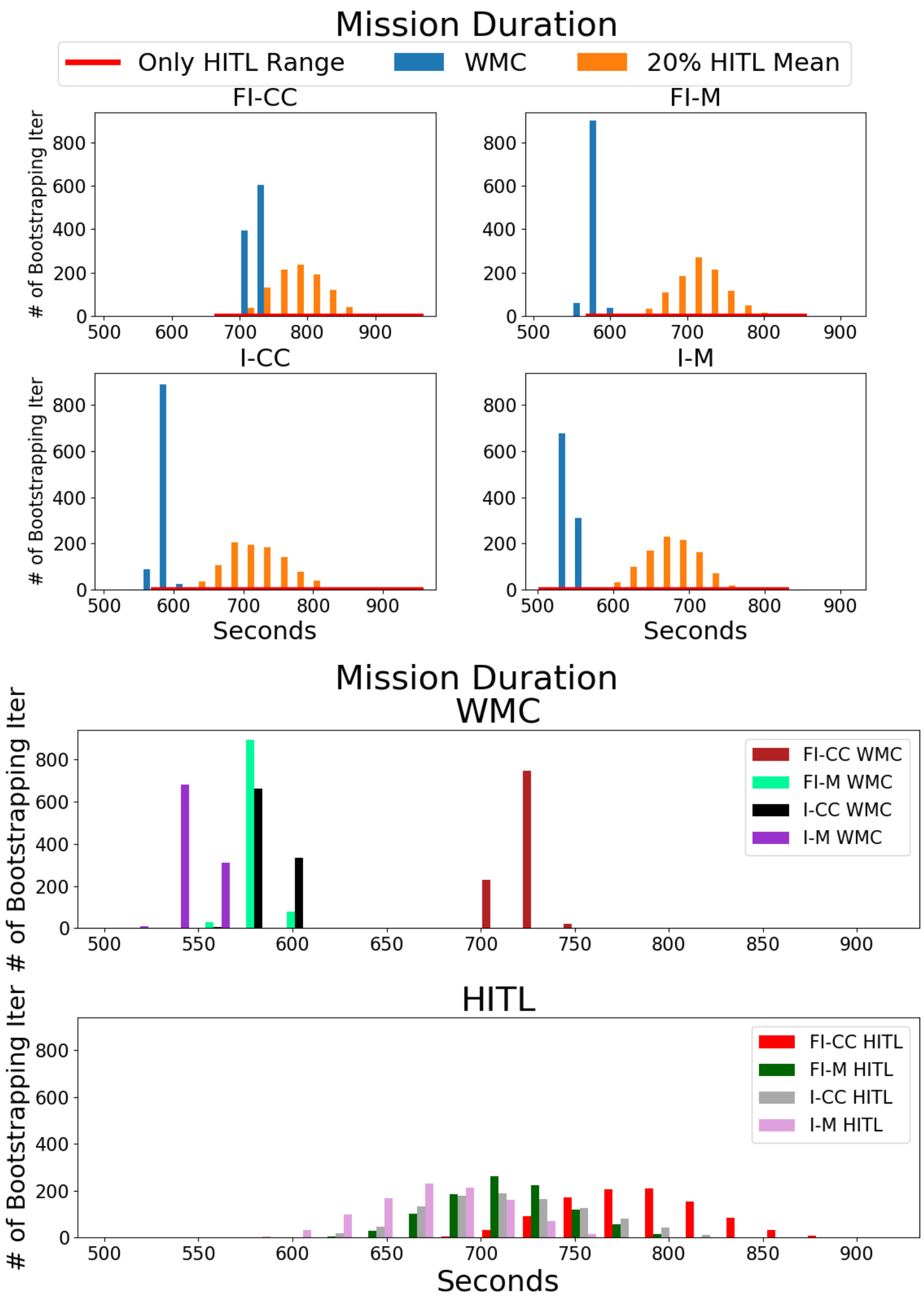


Figure 6.11: Histograms of mission duration as predicted by refined WMC simulation and measured in HITL

The refined WMC also predicted the same trend as found in the HITL experiment for total busy time across the team, see fig. 6.13. Here the lowest to highest busy time was I-M < I-CC < FI-CC < FI-M. Of note, total busy time was also the metric that WMC predicts the most accurately, with the range of WMC values falling within the second stage comparison with the remaining 20% of P.

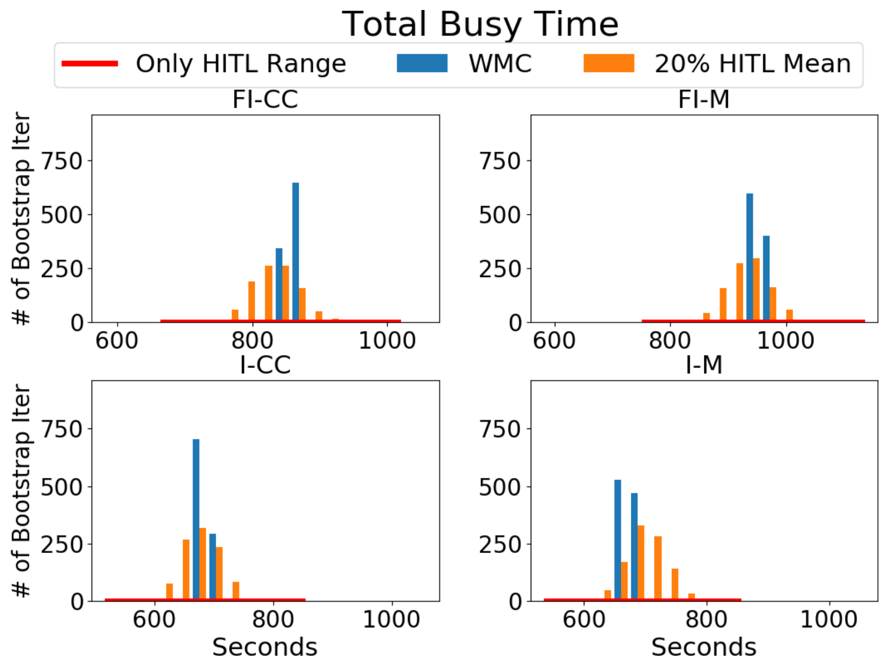


Figure 6.12: This graph shows a histogram of total busy time of each of the work allocations in separate quadrants. The blue columns represent the distribution of the WMC simulation data ($|W|$) while the orange columns represent the distribution of the averaged HITL data ($|V|$). The red line on the x-axis represents the range of the HITL data collected for that work allocation.

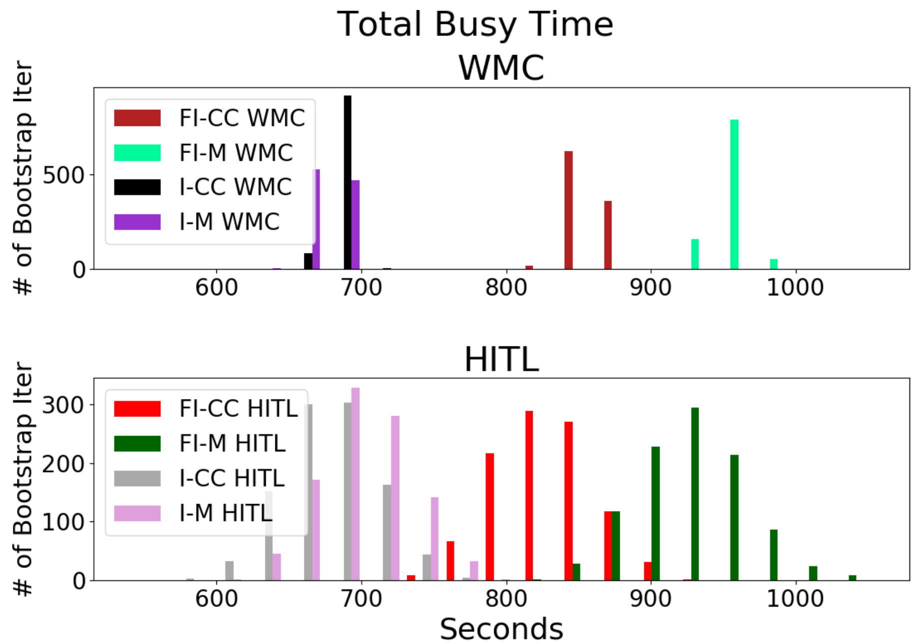


Figure 6.13: Histograms comparing total busy time metric grouped by WMC simulation and HITL data

For total idle time, the refined WMC predictions predicted the trend of work allocations from lowest to highest: FI-M < I-M < I-CC < FI-CC, see fig. 6.14. However, WMC underestimates this measure in its inherent assumptions that agents start actions immediately after completing another. Thus, relevant factors not been modeled in WMC (to date) include transition time between actions, and more communication actions than expected, i.e. that participants spoke with the IV before and after each action.

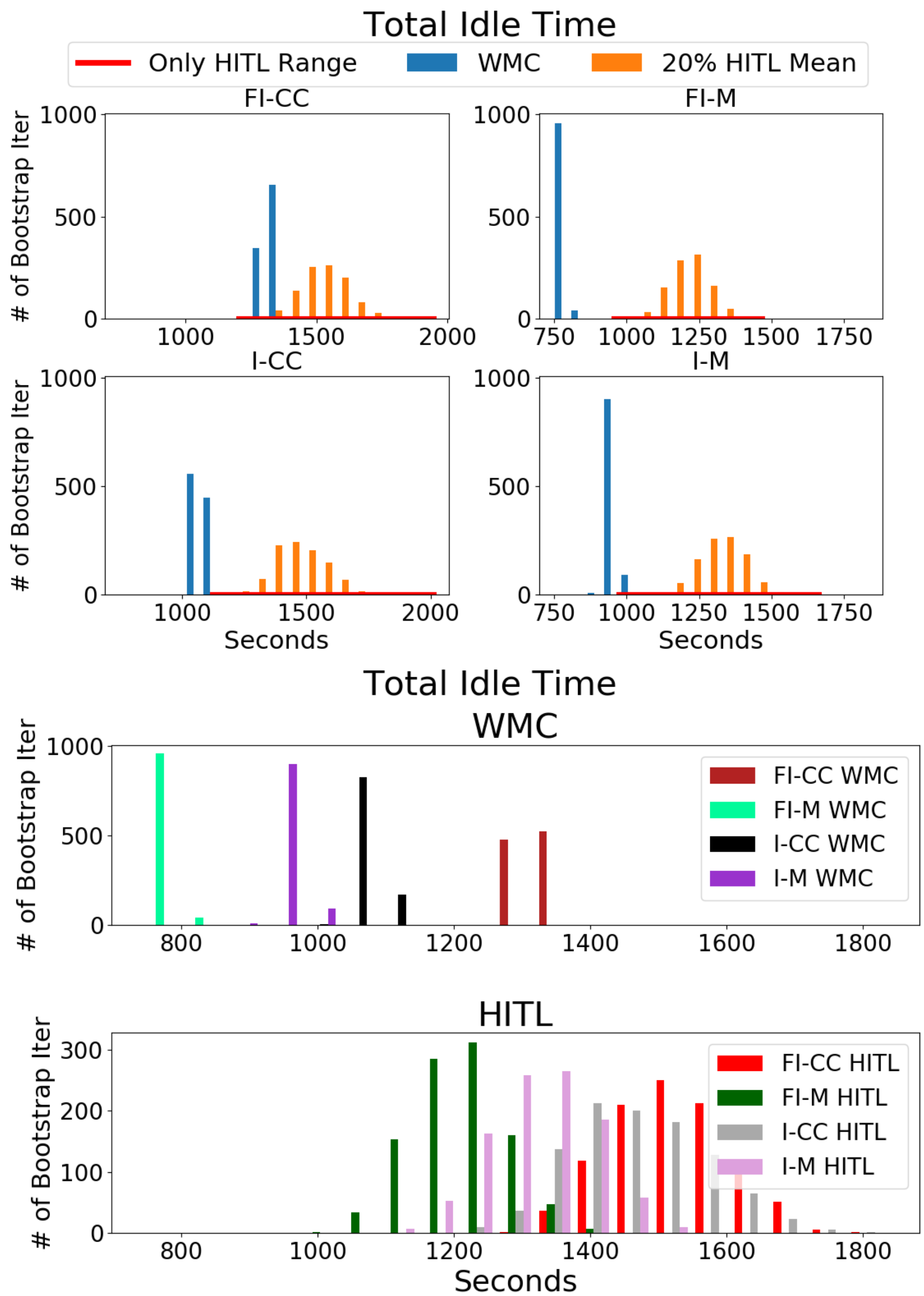


Figure 6.14: Histograms comparing total idle time metric grouped by WMC simulation and HITL data

Participant Metrics

In addition to the just-described assessments of time across the entire team, the experiment also assessed the timing of the participants' actions alone. Examining participant busy time, the refined WMC model can predict the trend between team conditions, but the HITL experiment results found greater variance than the WMC model predicts, see fig. 6.15. Similarly, the participants' idle time is accurate for trends between work allocation, see fig. 6.16. Finally, the time participants spent 'traversing', i.e. moving between stations, is well captured by the refined WMC results, see fig. 6.17.

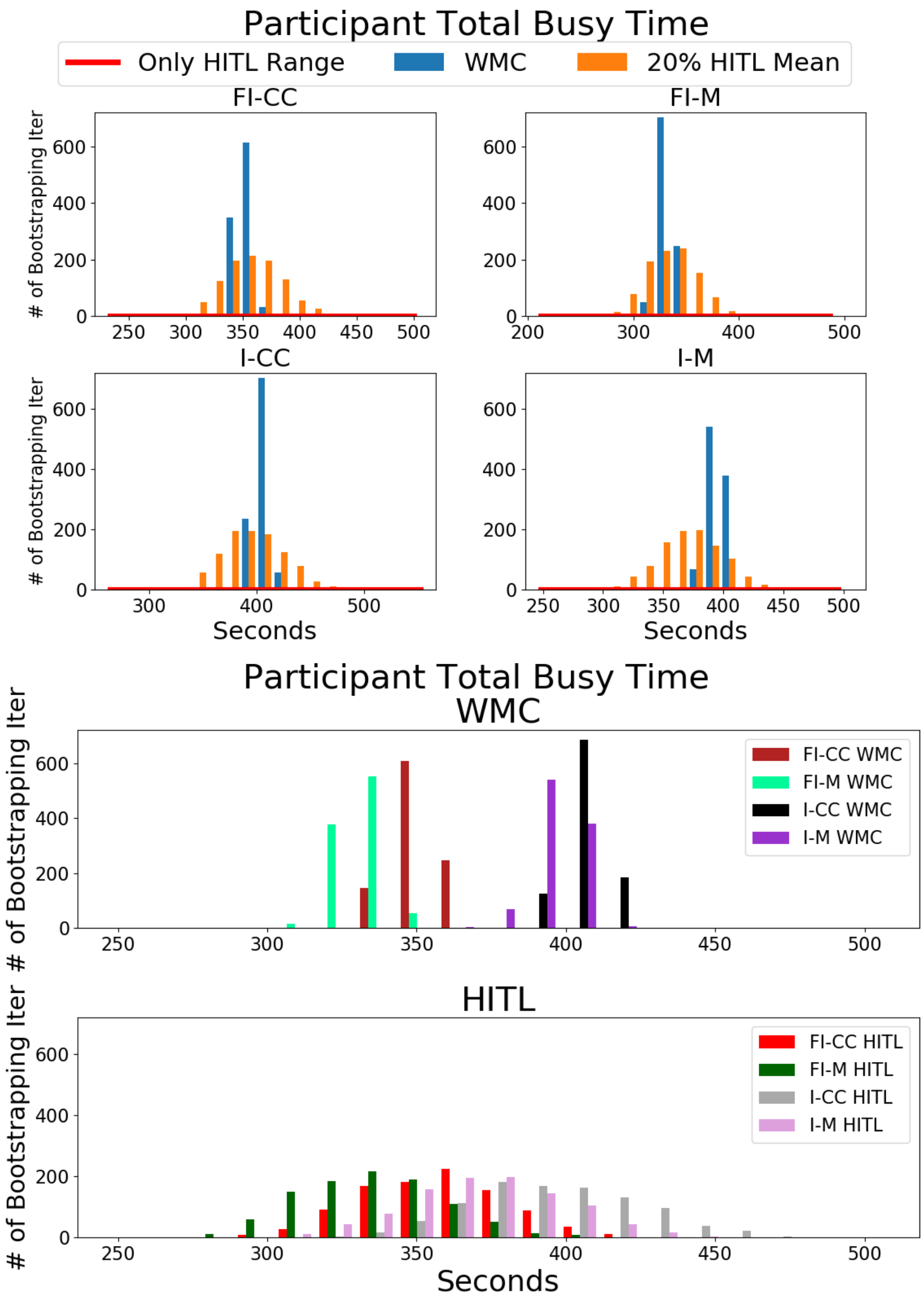


Figure 6.15: Histograms of participants' busy time as predicted by refined WMC simulation and measured in HITL. Same data is represented but grouped by work allocation (top) versus origin of data (bottom).

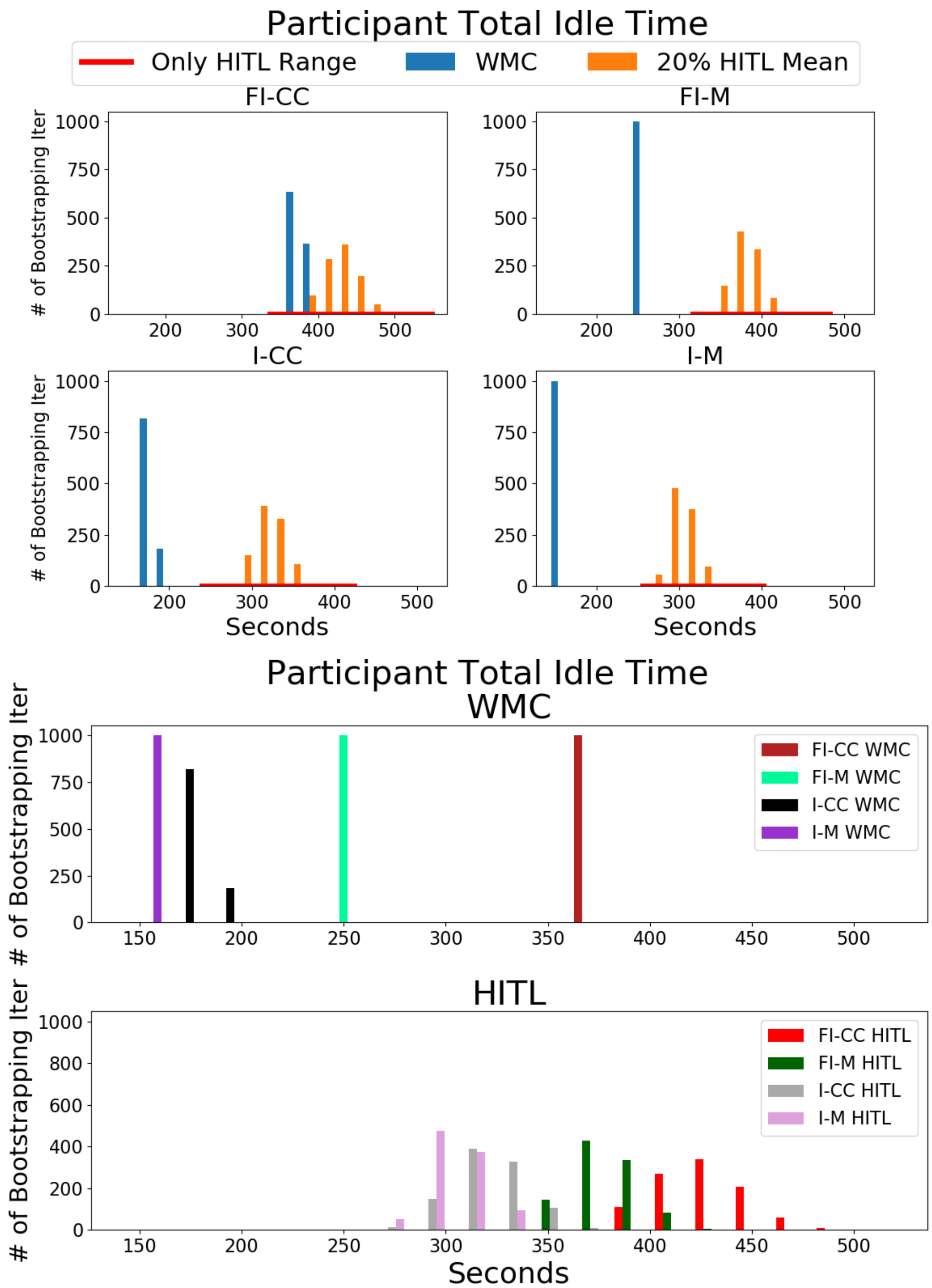


Figure 6.16: Histograms of participants' idle time as predicted by refined WMC simulation and measured in HITL

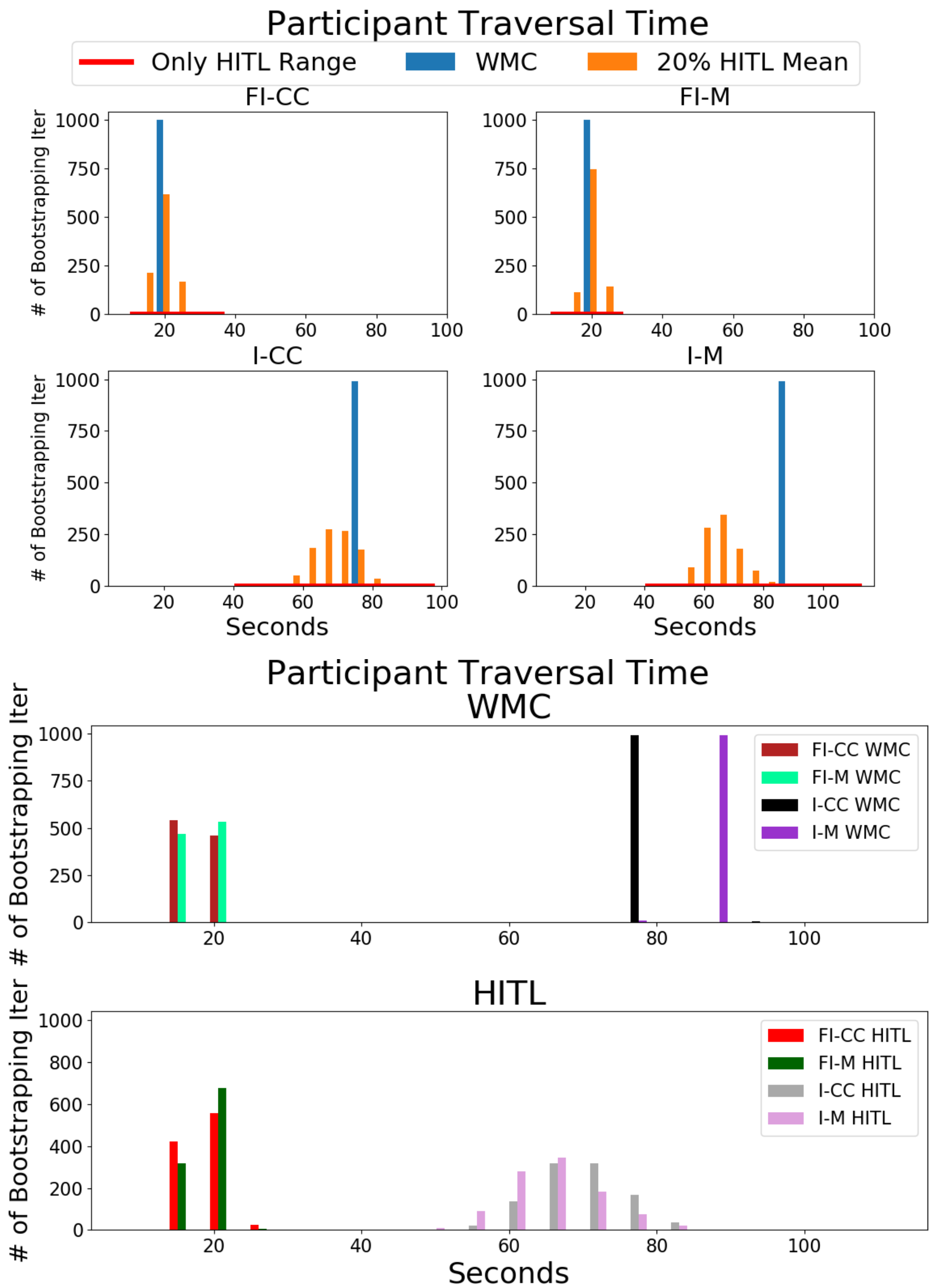


Figure 6.17: Histograms of participants' time spent traversing as predicted by refined WMC simulation and measured in HITL. Same data is represented but grouped by work allocation (top) versus origin of data (bottom).

6.4 Closing Thoughts

Design of a team, including defining task allocations and the teamwork modes used for human-robot interaction, can impact multiple measures of both the team and the work of individual team members. In this case, for example, the robot limited only to inspection tasks caused participants to expend more mental demand, physical demand, and effort. Changes in teamwork mode caused a similar effect on physical workload and effort, where robots that were commanded and confirmed (even both another teammate) required more attention from the participants. Overall, participants found the robotic agent with the additional fetching capabilities to be more effective as a teammate, regardless of tradeoff in a higher perceived idle time and requiring more attention. Thus, these types of design factors require careful modeling during the design of human-robot teams.

Through the analysis of this HITL, we can confirm our hypotheses that different work allocations and teamwork modes affect the participants' workload as well as how participants perceived the robotic agents. We can also see that changes in the capabilities of robotic agents, as well as the type of team interaction, can greatly impact an individual's perception of the robot and mission efficiency.

Computational models can predict the emergence and sequence of activity within the team, to an extent that can predict general trends within team design conditions. However, for these computational models to also provide accurate assessments of metrics of the time of activities, they need to be properly parameterized. This work demonstrated how the results of a HITL experiment can be applied by bootstrapping techniques to enable this parameterization. For measures of how much time the team members are busy at activities, and the resulting total mission duration, the refined computational model was found to accurately fit the HITL results. However, computational models to date have not captured the idle times where human agents are not busy, particularly the time spent communicating with teammates and the time spent transitioning between activities, and initiating new activities. These results suggest further improvements to computational models so that their comparative ability to quickly assess team designs can be used with greater accuracy and at all stages of team design.

CHAPTER 7
CONCLUSIONS AND FUTURE WORK

Human-Robot Teaming is a rich area with many facets to continue investigating. Though various research areas have dived deeply into teaming, there remains a lack of consensus on human-robot teaming concepts and evaluation methods. Team designers require tools to analyze and understand human-robot teaming to adjust to various scenarios, taskwork, and objectives.

Supporting the design of effective human-robot teaming is important to team designers and the team's success. As NASA is working towards future-day Mars and Lunar missions, understanding and supporting the integration of humans and robots in teams will be key to mission safety and success. Human-robot systems will need to be well integrated to successfully accomplish mission goals. This work supports team design by developing a series of methods and tools to investigate different teams early-in design through outer space case studies.

7.1 Contributions and Value Proposition

The central scope and theme of this work target the way we should design, evaluate, and think about human-robot teams. In doing so, I have defined a framework, conceptual methodology, and operationalized metrics for human-robot teams. In addition, this work has assessed various existing methodologies and perspectives to derive metrics operationalized from work allocation. Through these investigations, I have concluded that 1) interaction is the central construct in human-robot teaming and 2) the interaction can be greatly impacted by the definition of components in the teaming framework as well as the metrics used to characterize success. Designers and mission evaluators can use the framework alone or in conjunction with a simulation framework like WMC. In doing so we have addressed designer questions from Chapter 1 to help designers begin thinking about how to categorically define team requirements and understand which components to be aware of.

Afterward, we modeled and simulated human-robot teams using the WMC framework, specifically in an in-orbit maintenance scenario. Analysis of these simulated teams has implied the definition of an 'effective team' is dependent on the goals of the team designer. Understanding the metrics that matter to complete mission goals determines how teams should choose their metrics and framework components. This work, in particular, shows operationalized metrics at work. Additionally, the results help answer designer questions of how to take metrics, which there previously only defined in a qualitative manner, into consideration as well as provide a means to understand how to model and measure these metrics in a computational environment. Combining the HRT Framework and the WMC simulation framework can also predict median workload and metrics values/requirements of teams. Part of the value for designers then is allowing them to freely use simulation to evaluate teams and test different inputs to teams early on in design.

With HITL experimentation, we are further able to understand the human perspective of robotic team members as well as refine team modeling and validate our performance metrics. Overall, this work demonstrates the sensitivity of teams to both the definition and measurement of teamwork and taskwork components. Beyond demonstrating the usefulness of modeling and simulation to study human-robot teams, the HITL experiment paves the way for future real-life testing and analyses of these components within real human-robot teams.

7.2 Limitations of This Work

While we've shown many advantages of this work, we must also acknowledge the preliminary nature of this work. Limitations of this work must be understood in order to define better working conditions and settings through which to apply it.

Within the HRT Framework, there are a few key limitations that should be noted. First, not all human-human teaming components are specified in this framework, only those which apply to human-robot teaming have been included. For instance, human emotion, feelings, social norms, cues, and behaviors are unaccounted for. Therefore, to apply this framework to such teams would likely result in capturing only a portion of human-only teaming, as some aspects of these teams are ignored. As robots continue to advance, these elements may need to be included in the future. Second, as we analyze most teams on an individual basis, we have not explored the full application of this work for swarms. Our model and framework of the SWAT team in Chapter 2 serves as an initial observation of consideration of similar groups of individuals as one entity.

Within Work Models that Compute, WMC, our computational simulation framework, we surmise there are limitations in how the software was designed and implemented. Specifically, the current deterministic implementation of WMC, while purposefully implemented to help identify weakness that even perfect agent performance cannot rectify is greatly limiting. Many aspects of interaction and teaming are dynamic and unexpected. These have yet to be explored here. Additionally, while we have greatly expanded the modeling of teams and agent interaction in WMC, only a subset of team interaction and team composition are included in the latest version. We have yet to explore large teams and operations in WMC to date. WMC currently does not account for human limitations of variation although performance limitations are possible.

While the modeling and simulation of these teams can provide much insight into the metrics identified in this work, we cannot claim complete modeling of all aspects of human-robot teams defined in our framework. For one, external and random factors are largely unaccounted for, but modeling these in future work would provide a larger base of team performance to investigate.

7.3 Future Work

Future work for this HRT Framework could extend to specific frameworks that target different types of human-robot teams or configurations. There is much more work to be done in investigating archetypes of teams in human-robot teams as well as what types of behaviors and interactions to expect from these teams. Such teams could be identified by work patterns or robot types, such as groups of drones, or rover-type robots. Identifying more tailored frameworks for an ontology of teaming based on this larger foundation could be a possibility. Certainly, more depth and detail in each of the five components we identified is work to be done.

WMC could also be improved in various aspects, particularly with respect to the caveats in architecture design and detailed modeling of agents. While perfect agents prove a point in determining a conservative bound of expected behavior, there are many aspects of this to improve. For instance, a distinction between robotic agent models and human agent models could be so detailed as to define different types of robots or even humans. Human agents could also incorporate emotion, fatigue, and attention to start among other traits and abilities. Prior versions of the human agent model allowed for three internal lists of actions to be kept, active, interrupted and delayed. These agent models could be utilized to also mimic tasks that could be forgotten.

Additionally, future work could include modeling additional aspects of this framework (specifically, interactions) that we did not get to in this work so far. This could include more teamwork actions, dynamic teaming configurations, trust and fluency measures, internal cognition within agents, among many others mentioned throughout this work.

Beyond extending the HRT framework and computational simulation tool itself, there is much work to be done in using them to further investigate teams. First, future work could define and investigate a breadth of teams (since this work covers a depth analysis of one type of mission work). Second, further human-in-the-loop experiments with teams to test more interactions, mission scenarios, and metric results. This could include better modeling of cognitive processes time durations when communicating actions to one another, as pointed out by the HITL results. Third, much of what has been defined could be extrapolated to aspects of computational analysis that have yet to be explored here. One idea would be to create an optimization problem that defines this teaming problem. This problem could place weights on metrics that matter and provide information of which configurations of team members or taskwork allocation provides the best results given our metrics.

7.4 Utilizing This Work

We hope that aspects of this work can be applied in the real world. This work can be segmented and used together, separately, or paired with other systems.

Using the HRT Framework alone, the greatest takeaway is the ability to better understand the team in question given these five key factors. This work provides a thorough investigation of human-robot teaming to provide descriptions of not only the scope of human-robot teams but also which aspects to consider when developing a team. More specifically, the framework and methodology allow researchers to target interactions to focus on. Designers can use this framework to break down their team designs by identifying the five components of teaming in Chapter 2. Using the methodology in that chapter, designers can observe and define the aspects of teamwork that are important to their research questions, and also identify other components that are related and could be pursued further. This work also provides a compilation of distinct metrics and perspectives to measure teams to better aid designers in defining metrics that matter. This allows them to build these metrics directly into whatever system is being used to evaluate teams.

When using this baseline work in conjunction with WMC or another simulation framework, this work can provide insight into several factors for investigating team behavior and performance. Within WMC specifically, developers can increase modeling of any of the given components of teaming in more depth. For other simulations, developers can compare how subjective aspects of team components and team metrics can be defined in software. The modeling of these components in WMC then serves as a baseline to observe how team members, teamwork actions, and subjective metrics can be modeled and measured.

Overall, we hope that this work is able to provide answers to how designers can design, measure, and think about human-robot teams. Each aspect of this question is an entire space to dive deeply and this work tries to encapsulate a holistic top-level analysis of them. My hope is this research not only answers what is an important affecter in interaction for human-robot teams, but also why. This work is the initial legwork for investigating and compiling the breadth of factors that contribute to human-robot teaming, as well as evaluating said factors through modeling, simulation and data visualization methods. This work is only the beginning of discovering and categorizing key components of human-robot teams. Team interaction designers will be able to use this work to further understand and improve different team designs. Future work should continue to develop and expand these frameworks and models of human-robot teams.

Appendices

APPENDIX A

CHAPTER 5 APPENDIX

These appendices span the individual agent results for the two team configurations.

A.1 Team 1

Team 1 includes IV, EV, Fetching robot, and Humanoid agent.

A.1.1 Work Allocation 1 (FA1)

Total Times and Count Information

Total Mission Length: 5276.0

Total Task Duration Time: 11525

Total Busy Time: 9645

Total Idle Time: 6183.0

Total Taskload: 162

Total Traversal Count: 42

Total Fetch Count: 17

Total Failure Fixing Time: 0

The sub-teams for failures and the times they worked together are: []

Total Communication Information

Total Communication Instances: 215

Communication Start Time: 241.0

Communication End Time: 5171.0

Total Duration of Communication: 4930.0

Channel datalink2 was used : 75 times

Channel No_common_channel was used : 86 times

Channel datalink1 was used : 54 times

The agent pair ('IVastronaut', 'Humanoid') communicated a total of 75 times

The agent pair ('Humanoid', 'EVastronaut') communicated a total of 86 times

The agent pair ('EVastronaut', 'IVastronaut') communicated a total of 54 times

Total Physical Interaction Information

Total Number of Unique Physical Interactions : 8

The agent pair ('EVastronaut', 'Humanoid') physically interacted a total of 8 times

Teamwork Action Taskload Count

Total Confirmation Count: 0

Total Monitoring Count: 25

Total Control Count: 25

Total Command Count: 0

Total Joint Count: 8

Highest and Lowest Agents

Agent with highest task duration: ('EVastronaut', 4360)

Agent with lowest task duration: ('Humanoid', 3205)

Agent with highest idle time: ('IVastronaut', 3196.0)

Agent with lowest idle time: ('EVastronaut', 916.0)

Agent with highest taskload: ('Humanoid', 57)

Agent with lowest taskload: ('IVastronaut', 50)

Agent with highest traversal count: ('Humanoid', 32)

Agent with lowest traversal count: ('IVastronaut', 0)

Agent with highest fetch count: ('Humanoid', 17)

Agent with lowest fetch count: ('EVastronaut', 0)

IV Astronaut Individual Results *Total Times and Count Information*

Total Task Duration Time: 3960

Total Busy Time: 2080

Total Idle Time: 3196.0

Total Taskload: 50

Total Traversal Count: 0

Total Fetch Count: 0

Total Communication Information

IVastronaut communicated a total of 129 times

IVastronaut communicated with agent Humanoid over datalink2 a total of 75 time(s):

IVastronaut communicated with agent EVastronaut over datalink1 a total of 54 time(s):

Total Physical Interaction Information

IVastronaut interacted with other agents 0 times

Teamwork Action Taskload Count

Total Confirmation Count: 0

Total Monitoring Count: 25

Total Control Count: 25

Total Command Count: 0

Total Joint Count: 0

EV Astronaut Individual Results *Total Times and Count Information*

Total Task Duration Time: 4360

Total Busy Time: 4360

Total Idle Time: 916.0

Total Taskload: 55

Total Traversal Count: 10

Total Fetch Count: 0

Total Communication Information

EVastronaut communicated a total of 140 times

EVastronaut communicated with agent Humanoid over No_common_channel a total of 86 time(s):

EVastronaut communicated with agent IVastronaut over datalink1 a total of 54 time(s):

Total Physical Interaction Information

EVastronaut interacted with other agents 8 times

EVastronaut physically interacted with agent Humanoid using tool Panel2 a total of 2 time(s):

EVastronaut physically interacted with agent Humanoid using tool Panel3 a total of

2 time(s):

EVastronaut physically interacted with agent Humanoid using tool Filter1 a total of 2 time(s):

EVastronaut physically interacted with agent Humanoid using tool Filter2 a total of 2 time(s):

Teamwork Action Taskload Count

Total Confirmation Count: 0

Total Monitoring Count: 0

Total Control Count: 0

Total Command Count: 0

Total Joint Count: 8

Humanoid Robot Individual Results *Total Times and Count Information*

Total Task Duration Time: 3205

Total Busy Time: 3205

Total Idle Time: 2071.0

Total Taskload: 57

Total Traversal Count: 32

Total Fetch Count: 17

Total Communication Information

Humanoid communicated a total of 161 times

Humanoid communicated with agent IVastronaut over datalink2 a total of 75 time(s):

Humanoid communicated with agent EVastronaut over No_common_channel a total of 86 time(s):

Total Physical Interaction Information

Humanoid interacted with other agents 8 times

Humanoid physically interacted with agent EVastronaut using tool Panel2 a total of 2 time(s):

Humanoid physically interacted with agent EVastronaut using tool Panel3 a total of 2 time(s):

Humanoid physically interacted with agent EVastronaut using tool Filter1 a total of 2 time(s):

Humanoid physically interacted with agent EVastronaut using tool Filter2 a total of 2 time(s):

Teamwork Action Taskload Count

Total Confirmation Count: 0

Total Monitoring Count: 0

Total Control Count: 0

Total Command Count: 0

Total Joint Count: 0

A.1.2 Work Allocation 2 (FA 2)

Total Times and Count Information

Total Mission Length: 5841.0

Total Task Duration Time: 12895

Total Busy Time: 9745

Total Idle Time: 7778.0

Total Taskload: 169

Total Traversal Count: 36

Total Fetch Count: 17

Total Failure Fixing Time: 0

The sub-teams for failures and the times they worked together are: []

Total Communication Information

Total Communication Instances: 269

Communication Start Time: 241.0

Communication End Time: 5736.0

Total Duration of Communication: 5495.0

Channel datalink2 was used : 151 times

Channel No_common_channel was used : 91 times

Channel datalink1 was used : 27 times

The agent pair ('IVastronaut', 'Humanoid') communicated a total of 151 times

The agent pair ('Humanoid', 'EVastronaut') communicated a total of 91 times

The agent pair ('IVastronaut', 'EVastronaut') communicated a total of 27 times

Total Physical Interaction Information

Total Number of Unique Physical Interactions : 0

Teamwork Action Taskload Count

Total Confirmation Count: 0

Total Monitoring Count: 34

Total Control Count: 34

Total Command Count: 0

Total Joint Count: 3

Highest and Lowest Agents

Agent with highest task duration: ('IVastronaut', 6300)

Agent with lowest task duration: ('EVastronaut', 2220)

Agent with highest idle time: ('EVastronaut', 3621.0)

Agent with lowest idle time: ('Humanoid', 1466.0)

Agent with highest taskload: ('IVastronaut', 68)

Agent with lowest taskload: ('EVastronaut', 37)

Agent with highest traversal count: ('Humanoid', 30)

Agent with lowest traversal count: ('IVastronaut', 0)

Agent with highest fetch count: ('Humanoid', 17)

Agent with lowest fetch count: ('EVastronaut', 0)

IV Astronaut Individual Results *Total Times and Count Information*

Total Task Duration Time: 6300

Total Busy Time: 3150

Total Idle Time: 2691.0

Total Taskload: 68

Total Traversal Count: 0

Total Fetch Count: 0

Total Communication Information

IVastronaut communicated a total of 178 times

IVastronaut communicated with agent Humanoid over datalink2 a total of 151 time(s):

IVastronaut communicated with agent EVastronaut over datalink1 a total of 27 time(s):

Total Physical Interaction Information

IVastronaut interacted with other agents 0 times

Teamwork Action Taskload Count

Total Confirmation Count: 0

Total Monitoring Count: 34

Total Control Count: 34

Total Command Count: 0

Total Joint Count: 0

EV Astronaut Individual Results *Total Times and Count Information*

Total Task Duration Time: 2220

Total Busy Time: 2220

Total Idle Time: 3621.0

Total Taskload: 37

Total Traversal Count: 6

Total Fetch Count: 0

Total Communication Information

EVastronaut communicated a total of 118 times

EVastronaut communicated with agent Humanoid over No_common_channel a total of 91 time(s):

EVastronaut communicated with agent IVastronaut over datalink1 a total of 27 time(s):

Total Physical Interaction Information

EVastronaut interacted with other agents 0 times

Teamwork Action Taskload Count

Total Confirmation Count: 0

Total Monitoring Count: 0

Total Control Count: 0

Total Command Count: 0

Total Joint Count: 3

Humanoid Robot Individual Results *Total Times and Count Information*

Total Task Duration Time: 4375

Total Busy Time: 4375

Total Idle Time: 1466.0

Total Taskload: 64

Total Traversal Count: 30

Total Fetch Count: 17

Total Communication Information

Humanoid communicated a total of 242 times

Humanoid communicated with agent IVastronaut over datalink2 a total of 151 time(s):

Humanoid communicated with agent EVastronaut over No_common_channel a total of 91 time(s):

Total Physical Interaction Information

Humanoid interacted with other agents 0 times

Teamwork Action Taskload Count

Total Confirmation Count: 0

Total Monitoring Count: 0

Total Control Count: 0

Total Command Count: 0

Total Joint Count: 0

A.2 Team 2

Team 2 includes IV, EV, FFR, and RMS.

A.2.1 Work Allocation 1 (FA1)

Total Times and Count Information

Total Mission Length: 4109.1

Total Task Duration Time: 8594

Total Busy Time: 7695

Total Idle Time: 8741.400000000001

Total Taskload: 184

Total Traversal Count: 41

Total Fetch Count: 17

Total Failure Fixing Time: 0

The sub-teams for failures and the times they worked together are: []

Total Communication Information

Total Communication Instances: 259

Communication Start Time: 245.1

Communication End Time: 4004.1

Total Duration of Communication: 3759.0

Channel datalink2 was used : 105 times

Channel datalink1 was used : 154 times

The agent pair ('IVastronaut', 'Free_flying_robot') communicated a total of 119 times

The agent pair ('EVastronaut', 'Free_flying_robot') communicated a total of 98 times

The agent pair ('IVastronaut', 'RMS') communicated a total of 16 times

The agent pair ('Free_flying_robot', 'RMS') communicated a total of 14 times

The agent pair ('IVastronaut', 'EVastronaut') communicated a total of 12 times

Total Physical Interaction Information

Total Number of Unique Physical Interactions : 8

The agent pair ('EVastronaut', 'RMS') physically interacted a total of 8 times

Teamwork Action Taskload Count

Total Confirmation Count: 32

Total Monitoring Count: 8

Total Control Count: 8

Total Command Count: 33

Total Joint Count: 0

Highest and Lowest Agents

Agent with highest task duration: ('Free_flying_robot', 3175)

Agent with lowest task duration: ('RMS', 1035)

Agent with highest idle time: ('RMS', 3074.1000000000004)

Agent with lowest idle time: ('Free_flying_robot', 934.1000000000004)

Agent with highest taskload: ('IVastronaut', 81)

Agent with lowest taskload: ('RMS', 10)

Agent with highest traversal count: ('Free_flying_robot', 32)

Agent with lowest traversal count: ('IVastronaut', 0)

Agent with highest fetch count: ('Free_flying_robot', 17)

Agent with lowest fetch count: ('EVastronaut', 0)

IV Astronaut Individual Results *Total Times and Count Information*

Total Task Duration Time: 2084

Total Busy Time: 1185

Total Idle Time: 2924.1000000000004

Total Taskload: 81

Total Traversal Count: 0

Total Fetch Count: 0

Total Communication Information

IVastronaut communicated a total of 147 times

IVastronaut communicated with agent Free_flying_robot over datalink2 a total of 105 time(s):

IVastronaut communicated with agent Free_flying_robot over datalink1 a total of 14 time(s):

IVastronaut communicated with agent RMS over datalink1 a total of 16 time(s):

IVastronaut communicated with agent EVastronaut over datalink1 a total of 12 time(s):

Total Physical Interaction Information

IVastronaut interacted with other agents 0 times

Teamwork Action Taskload Count

Total Confirmation Count: 32

Total Monitoring Count: 8

Total Control Count: 8

Total Command Count: 33

Total Joint Count: 0

EV Astronaut Individual Results *Total Times and Count Information*

Total Task Duration Time: 2300

Total Busy Time: 2300

Total Idle Time: 1809.1000000000004

Total Taskload: 29

Total Traversal Count: 7

Total Fetch Count: 0

Total Communication Information

EVastronaut communicated a total of 110 times

EVastronaut communicated with agent Free_flying_robot over datalink1 a total of 98 time(s):

EVastronaut communicated with agent IVastronaut over datalink1 a total of 12 time(s):

Total Physical Interaction Information

EVastronaut interacted with other agents 8 times

EVastronaut physically interacted with agent RMS using tool Panel2 a total of 2 time(s):

EVastronaut physically interacted with agent RMS using tool Panel3 a total of 2 time(s):

EVastronaut physically interacted with agent RMS using tool Filter1 a total of 2 time(s):

EVastronaut physically interacted with agent RMS using tool Filter2 a total of 2 time(s):

Teamwork Action Taskload Count

Total Confirmation Count: 0

Total Monitoring Count: 0

Total Control Count: 0

Total Command Count: 0

Total Joint Count: 0

Free Flying Robot Individual Results *Total Times and Count Information*

Total Task Duration Time: 3175

Total Busy Time: 3175

Total Idle Time: 934.1000000000004

Total Taskload: 64

Total Traversal Count: 32

Total Fetch Count: 17

Total Communication Information

Free_flying_robot communicated a total of 231 times

Free_flying_robot communicated with agent IVastronaut over datalink2 a total of 105 time(s):

Free_flying_robot communicated with agent IVastronaut over datalink1 a total of 14 time(s):

Free_flying_robot communicated with agent EVastronaut over datalink1 a total of 98 time(s):

Free_flying_robot communicated with agent RMS over datalink1 a total of 14 time(s):

Total Physical Interaction Information

Free_flying_robot interacted with other agents 0 times

Teamwork Action Taskload Count

Total Confirmation Count: 0

Total Monitoring Count: 0

Total Control Count: 0

Total Command Count: 0

Total Joint Count: 0

Remote Manipulator System Individual Results *Total Times and Count Information*

Total Task Duration Time: 1035

Total Busy Time: 1035

Total Idle Time: 3074.1000000000004

Total Taskload: 10

Total Traversal Count: 2

Total Fetch Count: 0

Total Communication Information

RMS communicated a total of 30 times

RMS communicated with agent IVastronaut over datalink1 a total of 16 time(s):

RMS communicated with agent Free_flying_robot over datalink1 a total of 14 time(s):

Total Physical Interaction Information

RMS interacted with other agents 8 times

RMS physically interacted with agent EVastronaut using tool Panel2 a total of 2 time(s):

RMS physically interacted with agent EVastronaut using tool Panel3 a total of 2 time(s):

RMS physically interacted with agent EVastronaut using tool Filter1 a total of 2 time(s):

RMS physically interacted with agent EVastronaut using tool Filter2 a total of 2 time(s):

Teamwork Action Taskload Count

Total Confirmation Count: 0

Total Monitoring Count: 0

Total Control Count: 0

Total Command Count: 0

Total Joint Count: 0

A.2.2 Work Allocation 2 (FA 2)

Total Times and Count Information

Total Mission Length: 4951.0

Total Task Duration Time: 12280

Total Busy Time: 9520

Total Idle Time: 10284.0

Total Taskload: 185

Total Traversal Count: 43

Total Fetch Count: 17

Total Failure Fixing Time: 0

The sub-teams for failures and the times they worked together are: []

Total Communication Information

Total Communication Instances: 371

Communication Start Time: 241.0

Communication End Time: 4846.0

Total Duration of Communication: 4605.0

Channel datalink1 was used : 371 times

The agent pair ('EVastronaut', 'Free_flying_robot') communicated a total of 315 times

The agent pair ('IVastronaut', 'RMS') communicated a total of 16 times

The agent pair ('Free_flying_robot', 'RMS') communicated a total of 14 times

The agent pair ('Free_flying_robot', 'IVastronaut') communicated a total of 14 times

The agent pair ('IVastronaut', 'EVastronaut') communicated a total of 12 times

Total Physical Interaction Information

Total Number of Unique Physical Interactions : 8

The agent pair ('EVastronaut', 'RMS') physically interacted a total of 8 times

Teamwork Action Taskload Count

Total Confirmation Count: 0

Total Monitoring Count: 40

Total Control Count: 40

Total Command Count: 0

Total Joint Count: 0

Highest and Lowest Agents

Agent with highest task duration: ('EVastronaut', 6100)

Agent with lowest task duration: ('RMS', 1135)

Agent with highest idle time: ('IVastronaut', 3891.0)

Agent with lowest idle time: ('EVastronaut', 751.0)

Agent with highest taskload: ('EVastronaut', 93)

Agent with lowest taskload: ('RMS', 12)

Agent with highest traversal count: ('Free_flying_robot', 32)

Agent with lowest traversal count: ('IVastronaut', 0)

Agent with highest fetch count: ('Free_flying_robot', 17)

Agent with lowest fetch count: ('EVastronaut', 0)

IV Astronaut Individual Results *Total Times and Count Information*

Total Task Duration Time: 1920

Total Busy Time: 1060

Total Idle Time: 3891.0

Total Taskload: 16

Total Traversal Count: 0

Total Fetch Count: 0

Total Communication Information

IVastronaut communicated a total of 42 times

IVastronaut communicated with agent RMS over datalink1 a total of 16 time(s):

IVastronaut communicated with agent Free_flying_robot over datalink1 a total of 14 time(s):

IVastronaut communicated with agent EVastronaut over datalink1 a total of 12 time(s):

Total Physical Interaction Information

IVastronaut interacted with other agents 0 times

Teamwork Action Taskload Count

Total Confirmation Count: 0

Total Monitoring Count: 8

Total Control Count: 8

Total Command Count: 0

Total Joint Count: 0

EV Astronaut Individual Results *Total Times and Count Information*

Total Task Duration Time: 6100

Total Busy Time: 4200

Total Idle Time: 751.0

Total Taskload: 93

Total Traversal Count: 7

Total Fetch Count: 0

Total Communication Information

EVastronaut communicated a total of 327 times

EVastronaut communicated with agent Free_flying_robot over datalink1 a total of 315 time(s):

EVastronaut communicated with agent IVastronaut over datalink1 a total of 12 time(s):

Total Physical Interaction Information

EVastronaut interacted with other agents 8 times

EVastronaut physically interacted with agent RMS using tool Panel2 a total of 2 time(s):

EVastronaut physically interacted with agent RMS using tool Panel3 a total of 2 time(s):

EVastronaut physically interacted with agent RMS using tool Filter1 a total of 2 time(s):

EVastronaut physically interacted with agent RMS using tool Filter2 a total of 2 time(s):

Teamwork Action Taskload Count

Total Confirmation Count: 0

Total Monitoring Count: 32

Total Control Count: 32

Total Command Count: 0

Total Joint Count: 0

Free Flying Robot Individual Results *Total Times and Count Information*

Total Task Duration Time: 3125

Total Busy Time: 3125

Total Idle Time: 1826.0

Total Taskload: 64

Total Traversal Count: 32

Total Fetch Count: 17

Total Communication Information

Free_flying_robot communicated a total of 343 times

Free_flying_robot communicated with agent EVastronaut over datalink1 a total of 315 time(s):

Free_flying_robot communicated with agent RMS over datalink1 a total of 14 time(s):

Free_flying_robot communicated with agent IVastronaut over datalink1 a total of 14 time(s):

Total Physical Interaction Information

Free_flying_robot interacted with other agents 0 times

Teamwork Action Taskload Count

Total Confirmation Count: 0

Total Monitoring Count: 0

Total Control Count: 0

Total Command Count: 0

Total Joint Count: 0

Remote Manipulator System Individual Results *Total Times and Count Information*

Total Task Duration Time: 1135

Total Busy Time: 1135

Total Idle Time: 3816.0

Total Taskload: 12

Total Traversal Count: 4

Total Fetch Count: 0

Total Communication Information

RMS communicated a total of 30 times

RMS communicated with agent IVastronaut over datalink1 a total of 16 time(s):

RMS communicated with agent Free_flying_robot over datalink1 a total of 14 time(s):

Total Physical Interaction Information

RMS interacted with other agents 8 times

RMS physically interacted with agent EVastronaut using tool Panel2 a total of 2 time(s):

RMS physically interacted with agent EVastronaut using tool Panel3 a total of 2 time(s):

RMS physically interacted with agent EVastronaut using tool Filter1 a total of 2 time(s):

RMS physically interacted with agent EVastronaut using tool Filter2 a total of 2 time(s):

Teamwork Action Taskload Count

Total Confirmation Count: 0

Total Monitoring Count: 0

Total Control Count: 0

Total Command Count: 0

Total Joint Count: 0

REFERENCES

[1] Jessica J. Marquez et al. "Risk of Inadequate Design of Human and Automation / Robotic Integration Human Research Program Space Human Factors and Habitability Element". In: (2013), p. 42.

[2] Jessica J. Marquez et al. "Evaluation of Human and AutomationRobotics Integration Needs for Future Human Exploration Missions". In: (2016).

[3] Steve Kozlowski and Bradford Bell. "Work groups and teams in organizations: Review update". In: *Handbook of Psychology* 12 (2013), pp. 412–469. arXiv: `arXiv:1011.1669v3`.

[4] T Fong, C Thorpe, and C Baur. "Collaboration, dialogue and human-robot interaction, 10th international sumposium of robotics research (lorne, victoria, australia)". In: *Proceedings of the 10th International Symposium of Robotics Research*. 2001.

[5] Guy Hoffman and Cynthia Breazeal. "Collaboration in Human-Robot Teams". In: *Work* (2004), pp. 1–18.

[6] Felix Gervits et al. "Exploring Coordination in Human-Robot Teams in Space". In: *AIAA SPACE and Astronautics Forum and Exposition*. 2017, p. 5309.

[7] Steve W.J. J Kozlowski and Daniel R. Ilgen. "Enhancing the Effectiveness of Work Groups and Teams". In: *Psychological Science in the Public Interest* 7.3 (2006), pp. 77–124.

[8] John Mathieu et al. "Team effectiveness 1997-2007: A review of recent advancements and a glimpse into the future". In: *Journal of management* 34.3 (2008), pp. 410–476.

[9] Eduardo Salas, Dana E Sims, and C Shawn Burke. "Is there a "big five" in teamwork?" In: *Small group research* 36.5 (2005), pp. 555–599.

[10] Eric D Sundstrom et al. *Supporting work team effectiveness: Best management practices for fostering high performance*. Jossey-Bass Publishers San Francisco, 1999.

[11] Joseph E McGrath. "Toward a 'theory of method' for research on organizations". In: *New perspectives in organization research* 533 (1964), pp. 533–547.

[12] Daniel R Ilgen et al. “Teams in organizations: From input-process-output models to IMOI models”. In: *Annu. Rev. Psychol.* 56 (2005), pp. 517–543.

[13] Michelle A Marks, John E Mathieu, and Stephen J Zaccaro. “A temporally based framework and taxonomy of team processes”. In: *Academy of management review* 26.3 (2001), pp. 356–376.

[14] John Bradley, Barbara Jo White, and Brian E Mennecke. “Teams and tasks: A temporal framework for the effects of interpersonal interventions on team performance”. In: *Small Group Research* 34.3 (2003), pp. 353–387.

[15] Carsten KW De Dreu and Laurie R Weingart. “Task versus relationship conflict, team performance, and team member satisfaction: a meta-analysis.” In: *Journal of applied Psychology* 88.4 (2003), p. 741.

[16] Stanley M Gully et al. “A meta-analysis of team-efficacy, potency, and performance: interdependence and level of analysis as moderators of observed relationships.” In: *Journal of applied psychology* 87.5 (2002), p. 819.

[17] John E Mathieu, Lucy L Gilson, and Thomas M Ruddy. “Empowerment and team effectiveness: an empirical test of an integrated model.” In: *Journal of applied psychology* 91.1 (2006), p. 97.

[18] Dov Zohar. “A group-level model of safety climate: testing the effect of group climate on microaccidents in manufacturing jobs.” In: *Journal of applied psychology* 85.4 (2000), p. 587.

[19] Ad de Jong, Ko de Ruyter, and Jos Lemmink. “Service climate in self-managing teams: Mapping the linkage of team member perceptions and service performance outcomes in a business-to-business setting”. In: *Journal of Management Studies* 42.8 (2005), pp. 1593–1620.

[20] Jason A Colquitt, Raymond A Noe, and Christine L Jackson. “Justice in teams: Antecedents and consequences of procedural justice climate”. In: *Personnel psychology* 55.1 (2002), pp. 83–109.

[21] Daniel J Beal et al. “Cohesion and performance in groups: a meta-analytic clarification of construct relations.” In: *Journal of applied psychology* 88.6 (2003), p. 989.

[22] Roger C Mayer, James H Davis, and F David Schoorman. “An integrative model of organizational trust”. In: *Academy of management review* 20.3 (1995), pp. 709–734.

[23] J.E. Mathieu et al. “The influence of shared mental models on team process and performance.” In: *The Journal of applied psychology* 85.2 (2000), pp. 273–83.

[24] Franz W Kellermanns et al. “The lack of consensus about strategic consensus: Advancing theory and research”. In: *Journal of Management* 31.5 (2005), pp. 719–737.

[25] Susan G Cohen and Diane E Bailey. “What makes teams work: Group effectiveness research from the shop floor to the executive suite”. In: *Journal of management* 23.3 (1997), pp. 239–290.

[26] Eric Sundstrom et al. “Work groups: From the Hawthorne studies to work teams of the 1990s and beyond.” In: *Group Dynamics: Theory, Research, and Practice* 4.1 (2000), p. 44.

[27] Robert D Pritchard et al. “Effects of group feedback, goal setting, and incentives on organizational productivity.” In: *Journal of applied psychology* 73.2 (1988), p. 337.

[28] Prasad Balkundi and David A Harrison. “Ties, leaders, and time in teams: Strong inference about network structure’s effects on team viability and performance”. In: *Academy of Management Journal* 49.1 (2006), pp. 49–68.

[29] Herbert H Clark and Edward F Schaefer. “Contributing to discourse”. In: *Cognitive science* 13.2 (1989), pp. 259–294.

[30] Gareth R Jones and Jennifer M George. “The experience and evolution of trust: Implications for cooperation and teamwork”. In: *Academy of management review* 23.3 (1998), pp. 531–546.

[31] Michael A. Goodrich and Alan C. Schultz. “Human-Robot Interaction: A Survey”. In: *Foundations and Trends® in Human-Computer Interaction* 1.3 (2007), pp. 203–275. arXiv: `arXiv:1011.1669v3`.

[32] Robert R Hoffman and Laura G Militello. *Perspectives on cognitive task analysis: Historical origins and modern communities of practice*. Psychology Press, 2012.

[33] Karen M. Feigh and Amy R. Pritchett. “Requirements for Effective Function Allocation A Critical Review”. In: *Journal of Cognitive Engineering and Decision Making* 8.1 (2014), pp. 23–32.

[34] a. R. Pritchett, S. Y. Kim, and K. M. Feigh. “Modeling Human-Automation Function Allocation”. In: *Journal of Cognitive Engineering and Decision Making* 8.1 (2014), pp. 33–51.

[35] Julie Ann Shah, Joseph H. Saleh, and Jeffrey A. Hoffman. “Review and synthesis of considerations in architecting heterogeneous teams of humans and robots for optimal space exploration”. In: *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews* 37.5 (2007), pp. 779–793.

[36] Amy R Pritchett and So Young Kim. "Measuring Human-Automation Function Allocation". In: (2013).

[37] Sharon M Singer and David L Akin. "Role Definition and Task Allocation for a Cooperative EVA and Robotic Team". In: 4970 (2009).

[38] G. Rodriguez and C. R. Weisbin. "A new method to evaluate human-robot system performance". In: *Autonomous Robots* 14.2-3 (2003), pp. 165–178.

[39] Ayanna M. Howard. "A methodology to assess performance of human-robotic systems in achievement of collective tasks". In: *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS* (2005), pp. 377–382.

[40] Laurence Rognin, Pascal Salembier, and Moustapha Zouinar. "Cooperation, reliability of socio-technical systems and allocation of function". In: *International Journal of Human Computer Studies* 52 (2000), pp. 357–379.

[41] Raja Parasuraman, Thomas B. Sheridan, and Christopher D. Wickens. "A model for types and levels of human interaction with automation." In: *IEEE transactions on systems, man, and cybernetics. Part A, Systems and humans : a publication of the IEEE Systems, Man, and Cybernetics Society* 30.3 (2000), pp. 286–297.

[42] Kim J Vicente. *Cognitive work analysis: Toward safe, productive, and healthy computer-based work*. CRC Press, 1999.

[43] Mica R Endsley. *Designing for situation awareness: An approach to user-centered design*. CRC press, 2016.

[44] Ann M Bisantz and Catherine M Burns. *Applications of cognitive work analysis*. CRC Press, 2008.

[45] John Annett. "Hierarchical task analysis". In: *Handbook of cognitive task design*. CRC Press, 2003, pp. 41–60.

[46] Jan Maarten Schraagen, Susan F Chipman, and Valerie L Shalin. *Cognitive task analysis*. Psychology Press, 2000.

[47] Neville A Stanton et al. *Human factors methods: a practical guide for engineering and design*. CRC Press, 2017.

[48] Jens Rasmussen. "The role of hierarchical knowledge representation in decision-making and system management". In: *IEEE Transactions on systems, man, and cybernetics* 2 (1985), pp. 234–243.

[49] Jens Rasmussen. "Outlines of a hybrid model of the process plant operator". In: *Monitoring behavior and supervisory control*. Springer, 1976, pp. 371–383.

[50] Lanssie Mingyue Ma et al. "Human-Robot Teaming : Concepts and Components for Design". In: (), pp. 1–14.

[51] Martijn Ijtsma et al. "Modeling Human-Robot Interaction to Inform Function Allocation in Manned Spaceflight Operations". In: ().

[52] Martijn IJtsma et al. "Work Dynamics of Taskwork and Teamwork in Function Allocation for Manned Spaceflight Operations". In: *International Symposium on Aviation Psychology* (2017), pp. 554–559.

[53] Thomas B Sheridan and William L Verplank. *Human and computer control of undersea teleoperators*. Tech. rep. MASSACHUSETTS INST OF TECH CAMBRIDGE MAN-MACHINE SYSTEMS LAB, 1978.

[54] Stephen M Fiore. "Interdisciplinarity as Teamwork : How the Science of Teams Can Inform Team Science". In: *Small Group Research* 39.3 (2008), pp. 251–277.

[55] Andrea Thomaz and Guy Hoffman. "Computational Human-Robot Interaction". In: 4.2 (2016), pp. 105–223.

[56] Kerstin Dautenhahn. "Methodology & themes of human-robot interaction: A growing research field". In: *International Journal of Advanced Robotic Systems* 4.1 (2007), p. 15.

[57] Jean Scholtz. "Theory and evaluation of human robot interactions". In: *36th Annual Hawaii International Conference on System Sciences, 2003. Proceedings of the* 3 (2003), 10 pp.

[58] Gary Klein et al. "Common ground and coordination in joint activity". In: *Organizational simulation* 53 (2005), pp. 139–184.

[59] P R Cohen and H J Levesque. *Teamwork*. 1991.

[60] Matthew Johnson et al. *Coactive Design: Designing Support for Interdependence in Joint Activity*. Vol. 3. 1. 2014, p. 43. ISBN: 9789461863546.

[61] Raja Parasuraman et al. *Adaptive automation for human-robot teaming in future command and control systems*. Tech. rep. Army research lab aberdeen proving ground md human research and engineering, 2007.

[62] Ronald Wilcox, Stefanos Nikolaidis, and Julie Shah. "Optimization of temporal dynamics for adaptive human-robot interaction in assembly manufacturing". In: *Robotics* (2013), p. 441.

[63] Thomas Laengle, Thomas Hoeniger, and Lanjuan Zhu. "Cooperation in human-robot-teams". In: *ISIE'97 Proceeding of the IEEE International Symposium on Industrial Electronics*. IEEE. 1997, pp. 1297–1301.

[64] Claus Lenz and Alois Knoll. "Mechanisms and capabilities for human robot collaboration". In: *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*. IEEE. 2014, pp. 666–671.

[65] C Shawn Burke et al. "Understanding team adaptation: A conceptual analysis and model." In: *Journal of Applied Psychology* 91.6 (2006), p. 1189.

[66] Scott A Green et al. "Human-robot collaboration: A literature review and augmented reality approach in design". In: *International journal of advanced robotic systems* 5.1 (2008), p. 1.

[67] Stephen M Fiore et al. "Human-robot teams collaborating socially, organizationally, and culturally". In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 55. 1. SAGE Publications Sage CA: Los Angeles, CA. 2011, pp. 465–469.

[68] D Paul Benjamin, Deryle Lonsdale, and Damian Lyons. "A cognitive robotics approach to comprehending human language and behaviors". In: *2007 2nd ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE. 2007, pp. 185–192.

[69] Isabelle Blanchette and Anne Richards. "The influence of affect on higher level cognition: A review of research on interpretation, judgement, decision making and reasoning". In: *Cognition & Emotion* 24.4 (2010), pp. 561–595.

[70] Amos Freedy et al. "Measurement of trust in human-robot collaboration". In: *Proceedings of the 2007 International Symposium on Collaborative Technologies and Systems, CTS* (2007), pp. 106–114.

[71] Aaron Steinfeld et al. "Common Metrics for Human-Robot Interaction". In: *1st ACM SIGCHI/SIGART conference on Human-robot interaction (HRI '06)* (2006), pp. 33–40.

[72] Kevin Anthony Hoff and Masooda Bashir. "Trust in automation: Integrating empirical evidence on factors that influence trust". In: *Human Factors* 57.3 (2015), pp. 407–434.

[73] Christiano Castelfranchi and Rino Falcone. *Trust theory: A socio-cognitive and computational model*. Vol. 18. John Wiley & Sons, 2010.

[74] D Harrison McKnight and Norman L Chervany. “What trust means in e-commerce customer relationships: An interdisciplinary conceptual typology”. In: *International journal of electronic commerce* 6.2 (2001), pp. 35–59.

[75] Guy Hoffman. “Evaluating fluency in human-robot collaboration”. In: *International conference on human-robot interaction (HRI), workshop on human robot collaboration*. Vol. 381. 2013, pp. 1–8.

[76] Felix Gervits, Kathleen Eberhard, and Matthias Scheutz. “Team communication as a collaborative process”. In: *Frontiers in Robotics and AI* 3 (2016), p. 62.

[77] Sharolyn Converse, JA Cannon-Bowers, and E Salas. “Shared mental models in expert team decision making”. In: *Individual and group decision making: Current issues* 221 (1993).

[78] Susan Mohammed, Lori Ferzandi, and Katherine Hamilton. “Metaphor no more: A 15-year review of the team mental model construct”. In: *Journal of Management* 36.4 (2010), pp. 876–910.

[79] Renée J Stout et al. “Planning, shared mental models, and coordinated performance: An empirical link is established”. In: *Human Factors* 41.1 (1999), pp. 61–71.

[80] J Alberto Espinosa et al. “Shared Mental Models, Familiarity and Coordination: A Mulit-Method Study of Distributed Software Teams”. In: AIS. 2001.

[81] Catholijn M Jonker, M Birna Van Riemsdijk, and Bas Vermeulen. “Shared mental models”. In: *Coordination, organizations, institutions, and norms in agent systems vi*. Springer, 2011, pp. 132–151.

[82] Kyle B Reed and Michael A Peshkin. “Physical collaboration of human-human and human-robot teams”. In: *IEEE Transactions on Haptics* 1.2 (2008), pp. 108–120.

[83] Kevin T. Wynne and Joseph B. Lyons. “An integrative model of autonomous agent teammate-likeness”. In: *Theoretical Issues in Ergonomics Science* 19.3 (2018), pp. 353–374.

[84] Jean Scholtz et al. “Evaluation of human-robot interaction awareness in search and rescue”. In: *Robotics and Automation* (2004), pp. 2327–2332.

[85] Clint A Bowers, Curt C Braun, and Ben B Morgan Jr. "Team workload: Its meaning and measurement". In: *Team performance assessment and measurement*. Psychology Press, 1997, pp. 97–120.

[86] Holly A. Yanco, Jill L. Drury, and Jean Scholtz. "Beyond usability evaluation: Analysis of human robot interaction at a major robotics competition". In: *Journal of Human Computer Interaction* 19.November (2004), pp. 117–149.

[87] Sebastian Thrun. "Human ? Computer Interaction Toward a Framework for Human-Robot Interaction". In: 0024.June 2015 (2011), pp. 37–41.

[88] Jill L. Drury, Jean Scholtz, and H.a. Holly A. Yanco. "Awareness in human-robot interactions". In: *IEEE Systems, Man and Cybernetics* 1.October (2003), pp. 912–918.

[89] Ma Goodrich and Dr Olsen Jr. "Seven principles of efficient human robot interaction". In: *Systems, Man and Cybernetics, ...* 4 (2003), pp. 3942–3948.

[90] Matthew Johnson et al. "The fundamental principle of coactive design: Interdependence must shape autonomy". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 6541 LNAI (2011), pp. 172–191.

[91] Maryam Ashoori and Catherine M. Burns. "Team Cognitive Work Analysis Structure and Control Tasks". In: *Journal of Cognitive Engineering and Decision Making* 7.2 (2013), pp. 123–140.

[92] Maryam Ashoori. "Cognitive Work Analysis to Support Collaboration in Teamwork Environments". In: (2012), p. 242.

[93] Dan R Olsen and Michael A Goodrich. "Metrics for evaluating human-robot interactions". In: *Proceedings of PERMIS*. Vol. 2003. 2003, p. 4.

[94] Jean Scholtz. *Evaluation methods for human-system performance of intelligent systems*. Tech. rep. NATIONAL INST OF STANDARDS and TECHNOLOGY GAITHERSBURG MD MANUFACTURING, 2002.

[95] Lanssie Mingyue Ma et al. "Human-Robot Teaming: Concepts and Components for Design". In: *Field and Service Robotics*. Springer. 2018, pp. 649–663.

[96] Cynthia Breazeal et al. "Effects of nonverbal communication on efficiency and robustness in human-robot teamwork". In: *2005 IEEE/RSJ international conference on intelligent robots and systems*. IEEE. 2005, pp. 708–713.

[97] Bilge Mutlu et al. "Task structure and user attributes as elements of human-robot interaction design". In: *ROMAN 2006-The 15th IEEE International Symposium on Robot and Human Interactive Communication*. IEEE. 2006, pp. 74–79.

[98] T. Fong and I. Nourbakhsh. "Interaction challenges in human-robot space exploration". In: *Proceedings of the Fourth International Conference and Exposition on Robotics for Challenging Situations and Environments* January 2004 (2000), pp. 340–346.

[99] Laurel D Riek et al. "The Emerging Policy and Ethics of Human Robot Interaction." In: *HRI (Extended Abstracts)*. 2015, pp. 247–248.

[100] Aaron Steinfeld et al. "Common metrics for human-robot interaction". In: *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. ACM. 2006, pp. 33–40.

[101] Kartik Talamadupula et al. "Coordination in human-robot teams using mental modeling and plan recognition". In: *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE. 2014, pp. 2957–2962.

[102] David D Woods et al. "Envisioning human-robot coordination in future operations". In: *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 34.2 (2004), pp. 210–218.

[103] Sarah Woods et al. "Is this robot like me? Links between human and robot personality traits". In: *5th IEEE-RAS International Conference on Humanoid Robots, 2005.* IEEE. 2005, pp. 375–380.

[104] Felix Gervits, Terry Fong, and Matthias Scheutz. "Shared Mental Models to Support Distributed Human-Robot Teaming in Space". In: *2018 AIAA SPACE and Astronautics Forum and Exposition*. 2018, p. 5340.

[105] Peter A Hancock et al. "A meta-analysis of factors affecting trust in human-robot interaction". In: *Human Factors* 53.5 (2011), pp. 517–527.

[106] John D Lee and Katrina A See. "Trust in automation: Designing for appropriate reliance". In: *Human factors* 46.1 (2004), pp. 50–80.

[107] Robert M McIntyre and Eduardo Salas. "Measuring and managing for team performance: Emerging principles from complex environments". In: *Team effectiveness and decision making in organizations* (1995), pp. 9–45.

[108] Elizabeth Cha, Maja Matari, and Terrence Fong. "Nonverbal signaling for non-humanoid robots during human-robot collaboration". In: *ACM/IEEE International Conference on Human-Robot Interaction* 2016-April (2016), pp. 601–602.

[109] Terrence Fong et al. "The human-robot interaction operating system". In: *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. ACM. 2006, pp. 41–48.

[110] Maria Staudte and Matthew W Crocker. "Investigating joint attention mechanisms through spoken human–robot interaction". In: *Cognition* 120.2 (2011), pp. 268–291.

[111] Mica R Endsley and Daniel J Garland. *Situation awareness analysis and measurement*. CRC Press, 2000.

[112] Irving L Janis and Leon Mann. *Decision making: A psychological analysis of conflict, choice, and commitment.* free press, 1977.

[113] E De Visser et al. "A comprehensive methodology for assessing human-robot team performance for use in training and simulation". In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 50. 25. SAGE Publications Sage CA: Los Angeles, CA. 2006, pp. 2639–2643.

[114] Sarah Woods et al. "Comparing human robot interaction scenarios using live and video based methods: towards a novel methodological approach". In: *Advanced Motion Control, 2006. 9th IEEE International Workshop on*. IEEE. 2006, pp. 750–755.

[115] Jean Scholtz and Siavosh Bahrami. "Human-robot interaction: development of an evaluation methodology for the bystander role of interaction". In: *Systems, Man and Cybernetics, 2003. IEEE International Conference on*. Vol. 4. IEEE. 2003, pp. 3212–3217.

[116] Holly A Yanco, Jill L Drury, and Jean Scholtz. "Beyond usability evaluation: Analysis of human-robot interaction at a major robotics competition". In: *Human-Computer Interaction* 19.1 (2004), pp. 117–149.

[117] Stefanos Nikolaidis and Julie Shah. "Human-robot teaming using shared mental models". In: *ACM/IEEE HRI* (2012).

[118] Lanssie M Ma et al. "Modelling and Evaluating Failures in Human-Robot Teaming Using Simulation". In: (2018).

[119] Mark A Neerincx, Jurriaan van Diggelen, and Leo van Breda. "Interaction design patterns for adaptive human-agent-robot teamwork in high-risk domains". In: *International Conference on Engineering Psychology and Cognitive Ergonomics*. Springer. 2016, pp. 211–220.

[120] Tina Mioch et al. “Interaction design patterns for coherent and re-usable shape specifications of human-robot collaboration”. In: *Proceedings of the 2014 ACM SIGCHI symposium on Engineering interactive computing systems*. ACM. 2014, pp. 75–83.

[121] Hank Jones and Pamela Hinds. “Extreme work teams: using swat teams as a model for coordinating distributed robots”. In: *Proceedings of the 2002 ACM conference on Computer supported cooperative work*. ACM. 2002, pp. 372–381.

[122] Matthew Deans et al. “Robotic scouting for human exploration”. In: *AIAA Space 2009 Conference & Exposition*. 2009, p. 6781.

[123] Terrence Fong et al. “Field testing of utility robots for lunar surface operations”. In: *AIAA SPACE 2008 Conference & Exposition*. 2008, p. 7886.

[124] Terrence Fong et al. “Assessment of robotic recon for human exploration of the Moon”. In: *Acta Astronautica* 67.9-10 (2010), pp. 1176–1188.

[125] Maria Bualat et al. “Results from testing crew-controlled surface telerobotics on the International Space Station”. In: (2014).

[126] Becky L Hooey et al. “Modeling Operator Workload for the Resource Prospector Lunar Rover Mission”. In: *International Conference on Applied Human Factors and Ergonomics*. Springer. 2017, pp. 183–194.

[127] Robert Carvalho, Matthew Deans, and Jay Trimble. “Driving at the Lunar Poles: Simulations for Mission Design”. In: *2018 SpaceOps Conference*. 2018, p. 2471.

[128] Jay Trimble and Carvalho Robert. “Lunar prospecting: searching for volatiles at the south pole”. In: *14th International Conference on Space Operations*. 2016, p. 2482.

[129] Martijn Ijtsma, Amy R Pritchett, Raunak P Bhattacharyya, et al. “Computational simulation of authority-responsibility mismatches in air-ground function allocation”. In: (2015).

[130] Amy R Pritchett et al. “Simulating situated work”. In: *Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), 2011 IEEE First International Multi-Disciplinary Conference on*. IEEE. 2011, pp. 66–73.

[131] Amy R Pritchett. “22 Simulation to Assess Safety in Complex Work Environments”. In: *The Oxford handbook of cognitive engineering* (2013), p. 352.

[132] Lanssie M. Ma and Karen M Feigh. “Jumpstarting Modelling Systems Design”. In: *Human Factors and Ergonomics Society* (2017), pp. 718–722.

[133] Julie Ann Shah, Joseph H Saleh, and Jeffrey A Hoffman. “Review and synthesis of considerations in architecting heterogeneous teams of humans and robots for optimal space exploration”. In: *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 37.5 (2007), pp. 779–793.

[134] Tim Menzies et al. “Metrics that matter”. In: *Software Engineering Workshop, 2002. Proceedings. 27th Annual NASA Goddard/IEEE*. IEEE. 2002, pp. 51–57.

[135] Terrence Fong et al. “A preliminary study of peer-to-peer human-robot interaction”. In: *Systems, Man and Cybernetics, 2006. SMC’06. IEEE International Conference on*. Vol. 4. IEEE. 2006, pp. 3198–3203.

[136] Stefanos Nikolaidis et al. “Human-robot collaboration in manufacturing: Quantitative evaluation of predictable, convergent joint action”. In: *Robotics (isr), 2013 44th international symposium on*. IEEE. 2013, pp. 1–6.

[137] Stefanos Nikolaidis and Julie Shah. “Human-robot cross-training: computational formulation, modeling and evaluation of a human team training strategy”. In: *Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction*. IEEE Press. 2013, pp. 33–40.

[138] Przemyslaw A Lasota and Julie A Shah. “Analyzing the effects of human-aware motion planning on close-proximity human–robot collaboration”. In: *Human factors* 57.1 (2015), pp. 21–33.

[139] Przemyslaw A Lasota, Terrence Fong, Julie A Shah, et al. “A survey of methods for safe human-robot interaction”. In: *Foundations and Trends® in Robotics* 5.4 (2017), pp. 261–349.

[140] Matthew Gombolay et al. “Computational design of mixed-initiative human–robot teaming that considers human factors: situational awareness, workload, and workflow preferences”. In: *The International Journal of Robotics Research* 36.5-7 (2017), pp. 597–617.

[141] Elliot E Entin and Eileen B Entin. “Measures for evaluation of team processes and performance in experiments and exercises”. In: *Proceedings of the 6th International Command and Control research and Technology Symposium*. 2001, pp. 1–14.

[142] Jean MacMillan, Elliot E Entin, and Daniel Serfaty. “Communication overhead: The hidden cost of team cognition”. In: *Team cognition: Process and performance at the interand intra-individual level. American Psychological Association, Washington, DC.* (2004).

[143] Dirk Helbing and S. Balietti. *Agent-Based Modeling*. 2012, pp. 25–71. ISBN: 1853467960.

[144] Sybert H. Stroeve, Henk A.P. Blom, and G. J. Bakker. "Contrasting safety assessments of a runway incursion scenario: Event sequence analysis versus multi-agent dynamic risk modelling". In: *Reliability Engineering and System Safety* 109 (2013), pp. 133–149.

[145] J. Shah et al. "Analyzing air traffic management systems using agent-based modeling and simulation". In: *Proceedings of the 6th USA/Europe Air Traffic Management Research and Development* (2005), pp. 1–12.

[146] Sandra G Hart. "NASA-task load index (NASA-TLX); 20 years later". In: *Proceedings of the human factors and ergonomics society annual meeting*. Vol. 50. 9. Sage Publications Sage CA: Los Angeles, CA. 2006, pp. 904–908.