Title
The Imagined We: Shared Bayesian Theory of Mind for Modeling Communication

Author
Stacy, Stephanie Eu-Tien

Publication Date
2022

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

The Imagined We: Shared Bayesian Theory of Mind for Modeling Communication

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Statistics

by

Stephanie Eu-Tien Stacy

2022

ABSTRACT OF THE DISSERTATION

The Imagined We: Shared Bayesian Theory of Mind for Modeling Communication

by

Stephanie Eu-Tien Stacy
Doctor of Philosophy in Statistics
University of California, Los Angeles, 2022
Professor Tao Gao, Chair

How can a pointing gesture or simple utterance come to mean so much? Unlike a code or fixed mapping that can be used to uniquely identify a referent, humans flexibly use the same signal to mean many different things. As a result, to formally capture the impromptu, sparse nature of communication, it should be viewed as an inferential process. Both sending and understanding a signal in context requires reasoning about the underlying mind at the other end. This dissertation takes a cognitive approach to developing a computational framework for this type of uniquely human communication. Even young children who cannot yet speak in full sentences use simple gestures and utterances in uniquely flexible and intelligent ways. This highlights the promise of a reverse engineering approach: the underlying cognitive mechanisms and commonsense reasoning accumulated during pre-linguistic development become the foundation for modeling intelligent communication.

The modeling approach taken here formalizes this theoretical account by connecting and extending three lines of work which have traditionally been viewed as separate domains. First, I adopt an existing model that draws from game theory and probabilistic inference to formalize flexible signal understanding. Second, I integrate this with socially rational models

of individual agency, which involve understanding why individuals act the way they do in terms of their underlying mental states. This follows the tradition of viewing communication as an inference problem, where understanding a signal is about understanding the underlying mind that generated it. Here I argue communication can be viewed in terms of its use: signals are a special type of rational action that can be used to coordinate individuals. This perspective connects modeling flexible signaling to models of intentional agents. Finally, my last step is to shift from individual agency to treating communication as a cooperative, shared agency problem. While shared agency has been a promising approach for coordinating cooperators, it has not yet been modeled in conjunction with communication. This dissertation bridges this gap, leading to the development of a novel framework for communication, called the Imagined We (IW). I justify this cooperative shared agency approach by drawing from a wealth of behavioral evidence in developmental and comparative psychology demonstrating how communication can be viewed as a way to facilitate increasingly sophisticated cooperation. Through a set of simulations in cooperative tasks, I demonstrate theoretical advantages of this perspective. Moreover, I combine this with behavioral evidence that can begin to support some of the theoretical claims this model makes.

The dissertation of Stephanie Eu-Tien Stacy is approved.

Richard Alan Clarke Dale

Hongjing Lu

Yingnian Wu

Tao Gao, Committee Chair

University of California, Los Angeles

2022

To my mom and dad, for supporting me in every step

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I owe much of my success to the numerous people who have supported me at UCLA. To my advisor Tao Gao, I am inspired by the way you approach research by formalizing insights from many disciplines, and I have learned so much from working with you. More importantly, you have shown me what it is to be both a great mentor and a great scientist. It has also been a singular pleasure to watch the Visual Intelligence lab develop out of its infancy: in its earliest stages, it demonstrated the capacity for remarkably intelligent ideas and research which build a strong foundation for formalizing increasingly sophisticated concepts. Not only this, but it has embodied the topics my research focuses on: cooperation and communication. To the members of the Visual Intelligence Lab, it has been a pleasure to collaborate on research projects, present at conferences together, and discuss ideas as a We. Thank you especially to Lucy Zhao, who started as my first RA and continued on to be a peer and friend; Kaiwen Jiang — it's my pleasure to be your guinea pig; Aishni Parab, who has supported me both in research and in life; Victor Zhang, I admire your tech wizardry; and my RAs, Chenfei Li, Yiling Yun, Max Potter, and Siyi Gong — it has been so rewarding to work together and watch you develop into wonderful independent researchers.

I am also incredibly thankful for the mentors who have supported me in my journey to reach this point. To Mahtash Esfandiari, your teaching philosophy and the way you care about others shows and it has been a pleasure to work on statistical consulting at UCLA with you. To Safa Zaki, my undergraduate thesis advisor, thank you for helping me find my passion for research. And to Leon Phillips, my high school math teacher, thank you for teaching me how math is just a form of creative thinking, through logic puzzles and games.

To all of my friends both inside and outside my cohort who have accompanied me on the journey, thank you for creating memories I will cherish. Sam Baugh, Chris Strohmeier, Sydney Kahmann, and Ami Wulf, thank you for study sessions, crazy cooking experiments, and game nights. Allison Daly, Amara Thind, Eric Heinrichs, Ana Sias, Matt Miller, Mike

Lauria, and Camilla Marrero, thank you for beach volleyball, art parties, and adventures inside and outside of LA.

To Juan, thank you for embarking with me on the PhD journey and supporting me every step of the way. You are my half orange and push me to be the best version of myself.

Last and maybe most importantly, to my mom and dad, thank you for the unconditional love and support in all of my endeavours. I owe so much of my success to the time and effort you have both put into helping me grow.

Portions of this work were supported by a NSF Graduate Research Fellowship under Grant No. DGE-1650604 (2019-2022) and UCLA Dean's Scholar Award (2017-2019). Contributions to this manuscript are as follows. Chapter 2 and parts of the Introduction of this dissertation are adapted from a manuscript in prep written in conjunction with Siyi Gong, Aishni Parab, and Tao Gao. SG drafted parts focusing on developmental and comparative psychology and made editing contributions throughout the manuscript. AP contributed to edits on modeling cooperation, and TG contributed to theoretical components of the manuscript. Chapter 3 is expanded from Stacy, Parab, Kleiman-Weiner, & Gao (2022). TG and MKW gave valuable feedback in the development and contributed edits to this work. Chapter 4 is adapted from Stacy, Li, Zhao, Yun, Zhao, Kleiman-Weiner, & Gao (2021). TG made many theoretical contributions to the model and provided revision suggestions, CL produced modeling results for Simulation 3, and MKW provided thoughtful feedback on the manuscript and simulation experiments. Chapter 5 is adapted from Stacy, Yun, Potter, Moskowitz, & Gao (2022). YY and MP coded the experimental interface, helped with data analysis, and contributed to manuscript edits.

VITA

2013-2017	B.A. Statistics and Psychology (Honors), Williams College
2017-2019	Dean's Scholar, UCLA Graduate Division
2018-2022	Teaching Fellow, UCLA Department of Statistics
2019-2022	Graduate Research Fellow, National Science Foundation (NSF)
2021 Fall	Computer Vision/Artificial Intelligence Intern, General Electric Research

PUBLICATIONS AND PRESENTATIONS

Stacy, S., Yun, Y., Potter, M., Moskowitz, N., & Gao, T. (2022). No such thing as the average listener: Belief-driven versus action driven strategies in signaling. In *Proceedings of the 44th Annual Conference of the Cognitive Science Society.* Cognitive Science Society.

Stacy, S., Parab, A., Kleiman-Weiner, M., & Gao, T. (2022). Overloaded communication as paternalistic helping. In *Proceedings of the 44th Annual Conference of the Cognitive Science Society.* Cognitive Science Society.

Jiang, K., Dahmani, A., **Stacy, S.**, Jiang, B., Rossano, F., Zhu, Y., & Gao, T. (2022). What is the point? A Theory of Mind model of relevance. In *Proceedings of the 44th Annual Conference of the Cognitive Science Society.* Cognitive Science Society.

Stacy, S., Yun, Y., Potter, M., & Gao, T. (2022). Solving signaling ambiguity through cooperation. *Social Intelligence in Humans and Robots Workshop, Robotics: Science and Systems (RSS)*. Presented as spotlight talk.

Yun, Y., **Stacy, S.**, & Gao, T. (2021). The role of joint utility and pragmatic reasoning in cooperative communication. *Meaning in Context: Pragmatic Communication in Humans and Machines Workshop at 35th Conference on Neural Information Processing Systems*

Stacy, S., Li, C., Zhao, M., Yun, Y., Zhao, Q., Kleiman-Weiner, M., & Gao, T. (2021). Modeling communication to coordinate perspectives in cooperation. In *Proceedings of the 43rd Annual Conference of the Cognitive Science Society.* Cognitive Science Society.

Stacy, S., Li, C., Zhao, M., Kleiman-Weiner, M., & Gao, T. (2021). Modeling communication to coordinate perspectives in cooperation. *Social Intelligence in Humans and Robots Workshop, IEEE International Conference on Robotics and Automation (ICRA).*

Jiang, K., **Stacy, S.**, Wei, C., Chan, A., Rossano, F., Zhu, Y., & Gao, T. (2021). Individual vs. joint perception: A pragmatic model of pointing as communicative Smithian helping. In *Proceedings of the 43rd Annual Conference of the Cognitive Science Society.* Cognitive Science Society.

Karlin, J., Farajzadeh, J., **Stacy, S.**, Esfandiari, M., & Rootman, D. (2021). The effect of zero gravity on eyelid and brow position. *Ophthalmic Plastic and Reconstructive Surgery.*

Tang, N., **Stacy, S.**, Zhao, M., Marquez, G., Gao, T. (2020). Bootstrapping an Imagined We for cooperation. In *Proceedings of the 42nd Annual Conference of the Cognitive Science Society.* Cognitive Science Society.

CHAPTER 1

Introduction

Communication is a universal part of everyday life: we engage in it constantly, often without thought. For example, Audrey has plans to meet George at a restaurant for lunch. As she gets ready, she texts him that she is on her way. At the bus stop, she puts her hand out as she sees the bus approach. When she arrives, George is already sitting at a table. He raises an eyebrow and looks pointedly at his watch. Throughout this example, Audrey engages in a variety of communicative exchanges, all without saying a word. Communication is arguably one of the most important and sophisticated part of human intelligence. Not only can we communicate in different forms — through words, writing, gesture, and eye-gaze — but this communication is often ambiguous and contextual. Despite this, humans are extremely good at inferring rich meaning behind simple signals: when George looks at his watch, we infer Audrey is late. How can we say so much with so little? I approach this question by looking at communication in terms of its origin and development in humans. This dissertation proposes a computational account of communication that leverages cognitive insights to approach this type of human-unique overloaded and highly spontaneous communication.

1.1 A Language Approach to Modeling Communication

Without language, thought is a vague, uncharted nebula.

– Ferdinand de Saussure, Course in General Linguistics

Historically, language has been used as a critical benchmark to evaluate both human and machine intelligence. Early psychological theories emphasize language as arguably the most decisive characteristic contributing to unique human cognition and behavior. Starting from the late 1800's, language has been argued as a human instinct which enables flexible intelligence through interaction and competition with other instincts (Pinker, 2007 [1994]). Later, language was regarded as a means to shape reality through grammatical structure (e.g. Whorf, 1956) and as a biologically pre-programmed system laying the foundation for various cognitive capacities (Chomsky, 1983). Meanwhile, early intelligent machines have also emphasized language, through question answering. The well known Turing test evaluates artificial intelligence (AI) by distinguishing natural language responses from a human versus a machine (Turing, 2009 [1950]). From both fields, language generation and understanding has been treated as a cognitive divider to differentiate human-level intelligence. More broadly, it seems there is something intrinsically unique about language that makes it critical for achieving a deeper understanding of human intelligence. Certainly language in the context of syntax and grammar is exceedingly important in the study of more sophisticated forms of human communication; however, here, we emphasize a different modeling approach. This dissertation highlights how to formalize the underlying cognitive mechanisms necessary to reverse engineer human communication at its development.

This perspective comes in part from the recent body of evidence in comparative psychology which has shown that many of the types of intelligence that have been classically attributed to language are actually shared by other non-human species. For example, corvids exhibit complex behaviors in a variety of domains including the use and manufacture of tools (Emery & Clayton, 2004; Hunt, 1996). Perhaps the most noteworthy results come from studies on chimpanzees which are similar to humans in many respects, including elements of their capacity for sophisticated physical cognition such as tool use (Tomasello, Davis-Dasilva, CamaK, & Bard, 1987), working memory (Inoue & Matsuzawa, 2007), and object tracking (Barth & Call, 2006). More importantly, chimpanzees also have demonstrated their capacity for individual social cognition, which was originally thought to be human-unique. For example, chimpanzees understand what others can and cannot see (Hare, Call, & Tomasello, 2006) as well as others' goals (Warneken & Tomasello, 2006; Warneken, Hare, Melis, Hanus, & Tomasello, 2007; Yamamoto, Humle, & Tanaka, 2012) and intentions (Tomasello, Carpenter, Call, Behne, & Moll, 2005; Buttelmann, Carpenter, Call, & Tomasello, 2007). Strikingly, chimpanzees can even mentally simulate and physically manipulate others' perception and knowledge (Hare, Call, Agnetta, & Tomasello, 2000; Hare et al., 2006; Melis, Call, & Tomasello, 2006) and perform similar to infants in classic false-belief tasks (Krupenye, Kano, Hirata, Call, & Tomasello, 2016; Buttelmann, Buttelmann, Carpenter, Call, & Tomasello, 2017). Taken together, these studies indicate that many of the aspects thought to be human-unique *because* of language are, in fact, part of the foundation *for* language. They also point to a gap that must be bridged to achieve flexible human-like communication. What additional cognitive mechanisms are needed to explain the difference in how humans communicate?

1.2 A Codebook Approach to Modeling Communication

> *The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have meaning; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem.*
>
> – Claude Shannon, A Mathematical Theory of Communication

Classic linguistics has also posed two divergent — though not entirely mutually exclusive — theoretical approaches to modeling communication: a code model and an inferential model (Sperber & Wilson, 1986). The first, the code model, has been widely adopted in information theory and AI approaches to modeling communication. In traditional information theory, a codebook — containing a one-to-one mapping between a signal and its meaning — serves as context for communication. This formulation is primarily concerned with how to encode and decode a signal as well as how to design the codebook that minimizes the length of the code that needs to be transmitted. Uncertainty in transmission comes from signal degradation from noise in the communication channel: if that noise were removed, the code's meaning could be uniquely identified. In the past, people have proposed using a codebook to model human communication (Shannon, 1948; Lewis, 1969; Skyrms, 2010), but on its own, the code model is insufficient to explain spontaneous exchanges where the same signal can mean something completely different depending on the context (Sperber & Wilson, 1986; Misyak et al., 2016; Levinson, 1983). Many current AI approaches that do include communication treat it as a code which is assumed or learned (P. J. Gmytrasiewicz & Durfee, 2001; Oliehoek & Amato, 2016; Havrylov & Titov, 2017). This typically relies on a centralized planning mechanism or extensive training to learn how to map signals to referents or outcomes *a priori*. Critically, while these methods may develop language-like properties such as redundancy and hierarchy (Havrylov & Titov, 2017), they still fail at the human ability to use the same signal — a point, an eyebrow raise, an exclaimed "Hey!" — to mean a variety of things in context.

1.3 A Cognitive Approach to Modeling Communication

> *Point to a piece of paper. And now point to its shape — now to its color — now to its number ... How did you do it?*
>
> – Ludwig Wittgenstein, Philosophical Investigations

In contrast to the code model, the second classic linguistic approach to communication is the inferential model. Here signals are designed to provide evidence for one's mental states including beliefs, desires, and intentions; on the receiver side, understanding a signal requires inferring those mental states from that evidence (Sperber & Wilson, 1986; W. Levelt, 1989; Wharton, 2003). Instead of fixed meanings that can be encoded and decoded, meaning is flexible and comes from inferring the contents of the speaker's mind. Additionally, instead of treating uncertainty as noise in the signaling channel, uncertainty comes from inherent overloading in meaning, which must be resolved using information publicly and mutually known in the common ground.

The approach taken in this dissertation follows this inferential model and highlights the underlying cognitive mechanisms developed to support human-unique communication. In this way, commonsense knowledge accumulated during pre-linguistic development becomes part of the foundation for modeling intelligent communication. Specifically, our work formalizes this theoretical account by building upon and connecting three related but largely isolated modeling perspectives in the field.

1.3.1 Modeling Communication as a Rational Action

The first element of this approach is to use the Rational Speech Act (RSA) framework to model communication as an inferential process (Frank & Goodman, 2012; Goodman & Frank, 2016). This view is largely inspired by an influential account of pragmatics: the study of how signals can be interpreted in context. Grice (1975) proposed that communication should be taken as cooperative; communicators should be informative, truthful, relevant, and straightforward. While there have been many efforts to derive rules for pragmatics, the inferential view on communication has been notoriously difficult to formalize. Recently, Grice's initial observations on straightforward, cooperative communication have been re-examined from a computational perspective. The insight of the resulting RSA framework is to view communication as an optimization problem. With many signals to choose from, a speaker makes a selection that maximizes a utility function associated with those choices. Chapter 2.3 provides a more comprehensive description of how this utility function is defined and Chapters 3.1.4 and 3.2 extend and generalize this definition to bridge the gap between modeling perspectives. This frames communication as a classic decision theoretic problem where agents select actions rationally, according to their expected pay-offs. Here, signals serve as a particular type of rational action, and, as a result, inferring a signal's meaning is based on what kind of communicative intentions could have rationally produced that signal, giving a formal framework for sending and understanding overloaded communication.

1.3.2 Modeling Communication as Theory of Mind

The second aspect of this perspective is to view communicators as intentional agents with minds full of components such as beliefs, desires, and intentions. If RSA is an inferential model with signals as observations, this begs the question: what is the target of that inference? I argue that this target is the mind. Understanding the meaning of a signal in context requires the receiver to understand underlying informational, motivational, or intentional states of the speaker. This relies on a well-studied phenomena in psychology, Theory of Mind (ToM) which is the capacity to infer others' mental states based on their observed actions (Premack & Woodruff, 1978; Wellman, 1992).

As a result of intelligent reasoning under ToM, failing to correctly infer the contents of others' minds can cause problems in communication. A speaker who must explain every detail may blame the listener for being uninformed, incapable, or unmotivated. On the other hand, a listener may also be annoyed at a speaker who does not "get to the point" or "mansplains" aspects of a concept they are already well-acquainted with because the listener infers that the speaker does not hold their intelligence in high regard.

There have been many recent advances in cognition-based modeling that have taken a Bayesian approach to formalize ToM (Baker, Saxe, & Tenenbaum, 2009; Baker, Jara-Ettinger, Saxe, & Tenenbaum, 2017). However, while modeling communication under an inferential perspective is a ToM problem, this differs from classical ToM in two aspects. First, ToM is typically about understanding others' minds through their instrumental actions, which affect the environment directly. In communication, understanding others' minds occurs through their signals, which can only affect actions later downstream, *through* the mind. Second, modeling accounts of ToM focus on inferences made by an observer reasoning about an actor in the environment. In communication, agents are no longer observers but communicate and interact together. We link these concepts together formally in Chapter 3.

1.3.3 Modeling Communication through Shared Agency

The final aspect of my modeling approach takes the stance that communication, at its essence, should be understood in the context of cooperation: through humans' unique ability to collaborate as a group. Specifically, the intention of "working on something together as a 'We"' enables us to convey rich ideas with relatively sparse communication. This approach stems from a prominent theory that communication developed in the context of cooperation, as a social tool (Tomasello, 2010; Bruner, 1985; Vygotsky, 1978). From this standpoint, the purpose of communication is to facilitate increasingly sophisticated cooperation. When building a model, this suggests that a formalism of communication should start with a mechanism for cooperation.

In this sense, cooperation entails stronger constraints than being straightforward or informative as originally proposed by Grice. Here agents coordinate in mutually beneficial ways. This follows the tradition of viewing communication in terms of its use, as a type of joint action (Clark, 1996). At the same time, we are faced with instances of hostile communication both on a personal and international scale, which seems fundamentally at odds with a cooperative stance on communication. How then, can we make the argument that communication should start with a model of cooperation? I answer this question in the remainder of this section with evidence from both intuitive examples and rigorous empirical studies that can support and justify this modeling perspective.

First, much of the communication that is seen as non-cooperative still has important cooperative aspects. Interacting individuals hold many conflicting interests which make coordination and cooperation difficult, but it is precisely this conflict that requires sophisticated coordination through communication. Negotiation, bargaining, or even outright arguing is a form of this coordination — all parties continue to communicate in the hopes that they can gain, not relative to each other, but to their own personal values (Schelling, 1960). Communication allows different parties with different perspectives and interests to "get on the same page." It is not conflict in communication, but the absence of communication that marks a true departure from cooperation. For example, friends who argue are still friends, but friends who no longer want to talk to each other often drift apart. Between nations, an ambassador's first line of diplomacy is to negotiate, but recalling that ambassador to publicly close communication between countries acts as a signal of strong diplomatic censure (Regan, Frank, & Aydin, 2009). Moreover, establishing communication entails cooperative obligations: it can be considered rude to ignore someone talking to you, even if that person is a stranger. In these cases, communication does not even need to be verbal; for example, students in a classroom may avoid eye-contact when they do not want to be called on to avoid an obligation to participate that comes from meeting the professor's gaze. In conjunction, these everyday experiences suggest that there is room to negotiate mutually beneficial outcomes even in hostile communication.

From an ontogenetic perspective, the idea that cooperation is the foundation of communication is supported by evidence from developmental and comparative psychology, especially through studies that contrast the abilities of young children and chimpanzees. Despite the incredible achievements made by chimpanzees in terms of individual cognition and ToM, discussed previously, a gap between how apes and young human toddlers collaborate begins to emerge after around 2 to 3 years of development (Wobber, Herrmann, Hare, Wrangham, & Tomasello, 2014). It is not a coincidence that this is also the point at which children begin to communicate flexibly. For example, toddlers point in a variety of ways (Tomasello, Carpenter, & Liszkowski, 2007) and can even point to communicate about absent but mutually known entities (Liszkowski, Schäfer, Carpenter, & Tomasello, 2009). Toddlers can also interpret communicative signals based on the context of a joint task, reacting to an ambiguous request for help (Tomasello & Haberl, 2003). These same tendencies are not observed in chimpanzees, who struggle to even understand the helping intention behind a simple pointing gesture towards hidden food (Tomasello, Call, & Gluckman, 1997), unless presented in a competitive setting (Hare & Tomasello, 2004).

While chimpanzees do have the capacity for certain gestures which are relatively flexible (Call & Tomasello, 2020), their ability for sophisticated communication is highly limited due to their collaborative failures. Chimpanzee collaboration is marked by "group behavior in I-mode" (Tuomela, 2007). Often, the dominant individual simply takes all rewards after a collaborative effort, demotivating the subordinate individual in future attempts (Melis, Schneider, & Tomasello, 2011). When acting together, chimpanzees do not regulate their own or others' commitment (Warneken, Chen, & Tomasello, 2006), nor do they represent a shared goal with complementary, reversible roles (Tomasello, Carpenter, & Hobson, 2005). Instead chimpanzees seem to use each other as social tools to achieve greater individual rewards (Tomasello, 2019). This is in stark contrast to human children who remain jointly committed to tasks (Warneken et al., 2006; Vaish, Carpenter, & Tomasello, 2016) and treat collaborating agents as equal partners (Hamann, Warneken, & Tomasello, 2012; Warneken, Lohse, Melis, & Tomasello, 2011; Hamann, Warneken, Greenberg, & Tomasello, 2011).

In tasks where children do begin to successfully cooperate, they also start to communicate in sparse but sophisticated ways. In the rare cases where one collaborating child tries to take more rewards than is fair, a simple "Hey!" from their partner is readily understood and successfully regulates the greedy child's behavior (Warneken et al., 2011; Engelmann & Tomasello, 2019). One task that directly shows how communication can make cooperation more robust is a Stag Hunt paradigm, which has been studied extensively in game theory to demonstrate the challenge of social cooperation. Cooperating is risky, but also leads to higher rewards; whereas, acting individually is safe but less rewarding. In this task, young children — but not chimpanzees — can use minimal communication such as eye contact and joint attention to offset the risks of challenging coordination to collect larger rewards through successful cooperation (Duguid, Wyman, Bullinger, Herfurth-Majstorovic, & Tomasello, 2014; Siposova, Tomasello, & Carpenter, 2018).

Together, this evidence suggests that overloaded communication guides and facilitates cooperation by coordinating individuals' shared intentions. Moreover, adopting a shared agency perspective imposes strong assumptions on what is appropriate, relevant, and cooperative to communicate which can help resolve overloading in communication. There has been a recent body of computational work that has leveraged shared agency modeling to enable robust cooperation in non-communicative settings which is discussed in depth in Chapter 2.2 (Tang, Stacy, Zhao, Marquez, & Gao, 2020; Kleiman-Weiner, Ho, Austerweil, Littman, & Tenenbaum, 2016; Wu et al., 2021). This dissertation builds upon that perspective by modeling shared agency to support sparse, overloaded communication.

1.3.4 Integration of Perspectives

While I have argued for three key components to modeling communication from a cognitive approach, I make the additional claim that they are well posed to be integrated. While Chapters 3 and 4 of this dissertation provide the formalism for this integration, a few conceptual insights are worth noting. First, RSA is based on the idea of speech acts and the use of language: signals are treated as a type of rational action. This creates a strong tie to ToM reasoning. While communication is different from an instrumental action in many ways, it can still be reasoned over and subjected to a utility maximization process. Moreover, much of communication, especially early communication, occurs face-to-face in conjunction with instrumental actions. Second, shared agency is closely tied to ToM which provides an account of individual agency. While not the same, we would expect both types of agency to follow a similar overall structure. As a result ToM serves as the foundation for modeling shared agency where the target of inference is now a shared We mind instead of an individual's mind.

1.4 Outline of Successive Chapters

Taken together, this dissertation aims to formalize the cognitive perspective on communication, starting from the social infrastructure already in place when linguistic communication in children begins to develop. Chapter 2 places the work of this dissertation in the context of recent computational modeling progress and provides motivation for the modeling framework adopted in later chapters. Specifically, I review computational progress in the three areas which make up the modeling perspective taken: theory of mind (ToM), cooperation (shared agency), and communication (RSA). The remaining chapters of this dissertation provide both computational formalism and behavioral evidence to integrate these ideas.

I first build up to this idea through computational work in Chapter 3 which starts by bridging the gap between signals and actions to model highly contextual signals in a shared task. This starting point is important as communication rarely occurs in a vacuum. While classical accounts of communication have highlighted how signals are designed to change beliefs, from a coordination standpoint those beliefs are ultimately important because they are designed to engender certain downstream effects in our actions. Thus, sending and interpreting a signal depends not just on how it can change other's beliefs, but also how it affects actions. The shared task setting provides additional common ground context to help agents instantaneously coordinate on meaning. Here I demonstrate how integrating beliefs and actions as constraints to communication can begin to capture the one-shot, extremely flexible interpretation of overloaded signals in a humanlike manner. Details are summarized in Chapter 3 and published works Stacy, Parab, Kleiman-Weiner, and Gao (2022); Jiang et al. (2021, 2022).

After establishing the connection between signals and actions in the environment, I extend this model to adopt a shared agency perspective in Chapter 4. This model, which I call the Imagined We (IW), combines both cooperation and communication under the same cognitive framework, unifying these three modeling directions. Now agents reason over *joint* believes under a *shared* goal which is constrained by which *joint* actions are rational to take in the environment. This joint perspective also acts as another constraint to the meaning of a signal: a signaler should not ask her partner to do things when it is more efficient for her to do it herself while acting under the cooperative logic of shared agency. Details are summarized in Chapter 4 and published works Stacy et al. (2021); Tang et al. (2020).

Following the computational framework, in Chapter 5, I run a set of preliminary behavioral experiments to compliment the modeling work done in Chapter 4. Here we validate our modeling paradigm and provide behavioral evidence that humans can adopt strategies that take beliefs and actions into account when resolving overloading in communication. Moreover, we examine whether humans have a preference for one type of reasoning over the other when they are put in conflict. Details are summarized in Chapter 5 and published work Stacy, Yun, Potter, Moskowitz, and Gao (2022). Finally, Chapter 6 summarizes the findings of the preceding chapters, poses some important theoretical contributions of this work, and proposes directions for future work.

CHAPTER 2

Bayesian Theory of Mind for Cooperation and Communication

Even at the onset of the field of artificial intelligence (AI), scientists have been concerned with uncovering the nature of the mind by understanding the psychological mechanisms underlying fundamentally human capabilities such as reasoning and thinking (Newell, Shaw, & Simon, 1957; Simon & Newell, 1962). In fact, the initial goal of AI was to "find how to make machines use language, form abstractions and concepts, and solve problems now reserved for humans." (Russell, 2019, p. 15). While there has always been a deep connection between cognitive science and AI, this connection has been particularly fruitful in the last 20 years where cognitive modeling has focused increasingly on incorporating insights from the development of AI (Lake, Ullman, Tenenbaum, & Gershman, 2017; Zhu et al., 2020).

We describe how cognitive models of commonsense knowledge based on intuitive physics and psychology have been successfully applied to model the mind underlying social behaviors. At the core of this approach is Theory of Mind (ToM): a well-studied psychological phenomena which has been formalized using Bayesian models. ToM provides an agency-based account of individuals, allowing them to successfully understand and anticipate others' behavior in terms of their underlying mental states in a variety of tasks and environments. Moreover, ToM can naturally accommodate an intuitive understanding of the environment through a utility calculus defining the rewards and costs associated with behaviors. The power of this formulation stems from its ability to connect mental states to the environment and thus capture aspects of social and physical commonsense. On the social side, behaviors are the rational product of beliefs, desires, and intentions, and on the physical side, the actual actions that should be taken given those mental states depend on features of the environment.

Starting from ToM, we review two research directions that have been explored as a result: cooperation and communication. First, reformulating ToM as a shared agency problem has been shown as a promising avenue for understanding commonsense intelligence in cooperation both through empirical findings and computational work. This is because contextual cues under the constraint of mutual cooperation can help reduce ambiguity in how to achieve cooperation. We explore insights in modeling shared agency and provide a specific example in a multi-agent coordination setting emphasizing this advantage. In line with the second direction, we also review applications of ToM to tasks with communication. Treating signals as a particular type of rational action, the same inferential approach used to reason about underlying beliefs, desires, or intentions from observed behaviors can also be used to infer the meaning of a signal. A probabilistic approach is particularly promising as it can allow for uncertainty and overloading in a signal's meaning.

2.1 Modeling Individual Intentionality: Bayesian Theory of Mind

ToM refers to the ability to spontaneously attribute an action to the underlying mental states that produced it, including beliefs, desires, and intentions (Premack & Woodruff, 1978; Gopnik & Meltzoff, 1997; Dennett, 1987; Wellman, 1992). For example, you observe someone take a phone call in the library. This action can be interpreted as:

- The person thinks it is appropriate to take a call in the library (belief).
- This call is extremely important and urgent to the person (desire).
- The person is purposefully trying to annoy other people in the library (intention).

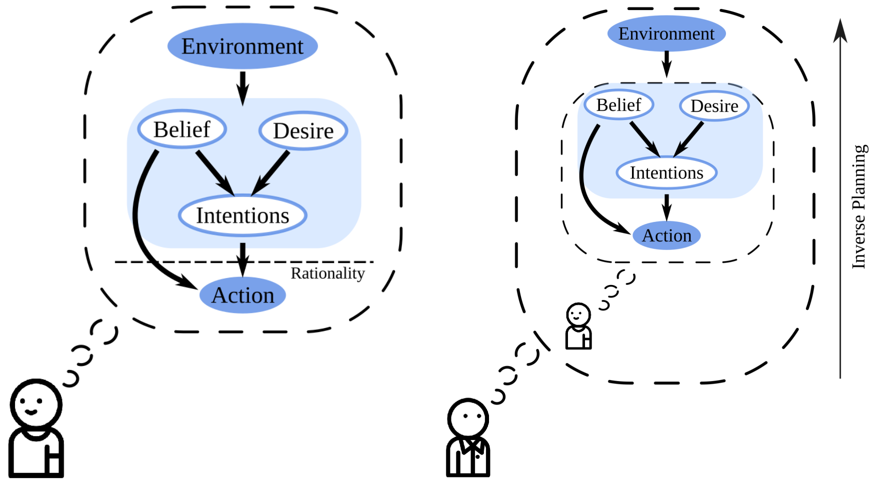
As a fundamental building block of social interaction, ToM plays a profound role in human society. For example, in our legal system a guilty act (*actus reus*) alone is often insufficient for conviction in many crimes: a guilty mind (*mens rea*) is also necessary. Moreover, *mens rea* is further divided into different levels of culpability which can be tied back to mental states: intent, knowledge, recklessness, and negligence. In addition to playing a huge role in society, ToM has been a topic of particular interest in the last several decades of development psychology (e.g. Wellman, 1992; Spelke & Kinzler, 2007; Gergely, Nádasdy, Csibra, & Bíró, 1995). Various studies have supported the early onset of such social capacity in humans, such as showing that infants as young as 6 months old can interpret desires and goals from agents' arm-reaching movement (Woodward, 1998).

Formally, ToM has been modeled through a Bayesian formulation: Bayesian Theory of Mind (BToM) (Baker et al., 2009). To interpret an action in terms of the observed mental states that generated it, BToM first requires a model of how actions are produced. This can be formalized through action planning, unfolded using a generative model. Forward planning uses the "principle of rationality" which asserts that an agent should act to maximize its expected utility while avoiding costs with respect to underlying mental states and constraints of the environment.

Actions are sampled from the soft-max of an agent's utility function, a method commonly used for approximately rational decision making (Luce, 1959), shown in Equation 3.4, where $\beta \in [0, \infty)$ describes the agent's degree of rationality. When $\beta = 0$, the agent is modeled as acting randomly, and as $\beta \to \infty$, the agent deterministically chooses the action with highest expected utility:

$$P(action|mind) \propto e^{\beta \mathbb{E}[U(action,mind)]} \tag{2.1}$$

In this framework, the mental states of others are made of three components: beliefs (b), desires (d), and intentions (i) following the tradition of Bratman (1987):



(a) Action Planning (b) Inverse Planning

Figure 2.1: 2.1a Rational action planning: In action selection, agents maximize their utility according to the mind. 2.1b Inverse planning: An observer sees another agent acting in the environment and infers their hidden mind.

$$P(mind) = P(b, d, i) = P(b)P(d)P(i|b, d) \tag{2.2}$$

Critically, using Bayesian inference, an observer can perform inverse planning to do backward reasoning over the generative models to infer the beliefs, desires, and intention of others that can best explain their observed actions (a) given the environment (w):

$$P(mind|a, w) \propto P(a|mind, w)P(mind|w) \tag{2.3}$$

This formulation was first tested by modeling human inference of an agent's goal in a 2D grid world (Baker et al., 2009). Thus, uncertainty in the *mind* in Equation 3.5 reduces to inferences about likely goals based on observed actions. The planning engine is implemented as a Markov Decision Process (MDP), which provides the probability of an action in a state conditioning on each possible goal. This model of goal inference has also been extended to adapt to changing or complex goals made of multiple sub-goals. However, even in its simplest form, by observing only a part of an agent's trajectory, the model can successfully infer the goal of that agent as well as capture temporal dynamics of human inferences in this task.

In follow-up studies, BToM has been extended to partially observable environments to jointly infer beliefs and desires (Baker & Tenenbaum, 2014; Baker et al., 2017). This demonstrates that rationally interpreting an agent's actions may require simultaneously reasoning over multiple components of the mind: here, the agent's uncertainty about the environment as well as personal preferences which determine how rewarding a particular goal is. Remarkably, employing BToM can sometimes even allow agents to make inferences about an agent's preference for a goal that is not currently in the environment

BToM has also been employed to formally capture human interpretation of intentionality purely from motion. For example, it can be used to infer perception of intentional goals such as exploration, attacking, gathering, and fleeing just from the physical dynamics of moving geometric shapes (Pantelis et al., 2014). In addition, BToM has also shown promising results capturing intentionality from motion in modeling the psychophysics of chasing under limited memory and attention (Gao, Baker, Tang, Xu, & Tenenbaum, 2019). Adding these constraints to BToM captures human performance and limitations when detecting predator-prey relations between pairs of targets in various conditions.

2.2 Extending BToM for Cooperation

ToM provides a scaffold for inference and planning and is often used to model the dynamics between an actor in the environment and a disconnected observer trying to understand the mind of that actor. However, cooperation requires individuals to simultaneously be both actors *and* observers. Thus, to capture the dynamic interactions in cooperation, agents need the capacity to perform both inverse planning — to understand one's partner — and forward planning — to generate one's own cooperative actions. Moreover, while ToM agents model their partner's mind from an individual perspective, such that "my mind understands your mind," cooperation often involves a joint perspective where individuals influence each other in a "meeting of minds." For this reason, our focus in this section is on one branch of modeling which has extended the individual perspective of classical ToM to incorporate shared agency or shared intentionality.

Philosophers have argued that the robust motivation for cooperation in humans is cognitively rooted in shared intentionality (Searle, 1990; Gilbert, 1992; Bratman, 1992). Under a shared intentionality framework, individuals intend that the self and others as a plural subject "We" — with its own actions and mind — jointly commit to a shared goal (Gilbert, 2013; Tomasello, 2010). This joint commitment is established through "readiness" of collaborators, must be normatively followed by all individuals, and cannot be rescinded unilaterally (Gilbert, 2013).

One challenge of modeling shared agency is its paradoxical nature: a shared intention emphasizes a collective representation of the group but must be implemented at an individual level (Schweikard & Schmid, 2021). Philosophical discourse posits that there are two core principles of shared intentionality. First, it is irreducible: shared intentionality is not simply the summation or aggregation of individual intentions, but rather has a qualitatively different structure. Second, it is individually owned: shared intentionality can only be experienced by an individual, within one's mind; thus, there is no ways to unilaterally dictate its formation, as forming a shared intention is voluntary. These two claims give rise to a tension: How can a collective intention that goes beyond any single individual simultaneously only make sense at the individual level?

Philosophers reconcile this paradox from different angles. One account argues that a We intention does not exist in reality but is instead imagined by each individual. For example, Margaret Gilbert focuses on the joint commitment, proposing that a shared intention is only realized when two or more individuals are willing to be "jointly committed to espousing a goal as a body" or as a "plural subject" (Gilbert, 2013, p. 30, 37). In parallel, Michael Bratman highlights the coordination of individuals' plans, proposing that a shared intention is an intricate mesh of individual plan-embedded intentions and their interrelations, aligned with each other and commonly known to all (Bratman, 1992, 2013). Importantly, shared intentionality as an infrastructure for cooperation also provides a new perspective on the nature of human-unique communication. That is, the purpose of communication is to coordinate agents' minds in a joint task so that each individual's distinct version of We can synchronize. While computational works following the tradition of both philosophical ideas are represented here, the modeling perspective in this dissertation most closely aligns with that of Gilbert.

Computational works modeling shared agency began to emerge in the 1990's and typically relied on logical language. Largely grounded in philosophy, Grosz and Kraus (1996) extended individual plans using a "sharedPlans" framework which contained both individual and mutual beliefs about how actions should be done. On the other hand, a model developed by Levesque, Cohen, and Nunes (1990) characterized the goal and intention of a group as "acting like a single agent" by defining a joint persistent goal that replaced individual beliefs in a personal goal with mutual belief. However, early shared agency formulations remained largely rule-based and disconnected from more recent Bayesian approaches for modeling ToM which represent uncertainty probabilistically instead of relying on rules.

More recently, probabilistic accounts of shared agency modeling based on BToM have begun to gain traction. They leverage Bayesian inference's capacity to reason over counterfactuals to answer (a) how shared intentions can be achieved when they are not real and (b) how different shared intentions imagined by individuals can be aligned during cooperation. The first instance of this modeling approach described a grid world coordination game where partners must infer whether they are competitors or cooperators and plan their actions accordingly (Kleiman-Weiner et al., 2016). The inference process to decide whether their partner is cooperative involves reasoning about whether the history of interactions can be explained by cooperative joint planning. During cooperation, shared intentions can be viewed as a form of augmented individual planning that takes the form "I intend that we J." Formally, this statement corresponds to agent i's generation of a *joint* policy π with respect to the goal G for each state s which makes action predictions for both herself (a_i) and her partner a_j. However, because agent i can only take her own action, the individual policy involves marginalizing out the partner: $\pi_i^G(s, a_i) = \sum_{a_j} \pi^G(s, a_i, a_j)$. Thus, different individuals are able to form an intention not only about one's own action but also about the states other agents reach in a "meshing of plans" (Bratman, 2013). This type of joint planning has also been proposed generically as a potential mechanism to plan and predict cooperative behavior in modeling work which aims to infer more complex team structures involving multiple levels of cooperative and competitive behavior (e.g. A is cooperating with B against C) (Shum, Kleiman-Weiner, Littman, & Tenenbaum, 2019).

In the formulation of Kleiman-Weiner et al. (2016), and most models of cooperation, a shared reward function determines how agents should act together. This joint BToM approach has also been applied to a collaborative cooking task as a means to coordinate while working together on a sub-task (e.g. chop the tomato) within a recipe (e.g. make a salad) (Wu et al., 2021). Throughout the task, agents work either in parallel on different sub-tasks or collectively coordinate on the same step to complete a recipe. Collaboration on a specific step may be necessary (e.g. if a counter divides the kitchen space then agents may need to pass ingredients across) or simply more efficient. This work captures the changing and reactive dynamics of coordinating under a shared task, using a joint form of ToM reasoning in order to do so; at the same time, it also introduces more complex task and environment elements.

2.2.1 Coordinating under the Imagined We framework

This dissertation adopts a stronger notion of cooperation that entails Gilbert's notion of cooperating "as a body." We have started by modeling BToM shared agency from a joint commitment perspective (Tang et al., 2020). To form a shared intention, each agent individually demonstrates a readiness to commit to the joint goal through her behavior which, in this case, does not require explicit communication. To achieve this, collaborators view themselves not as individuals, but rather as part of an imagined, supra-individual group We entity acting under the shared beliefs, shared desires, and shared intentions of the group (Gilbert, 1999). This focuses on cooperation beyond goal or reward sharing: namely that it entails a joint commitment among cooperators and the plural subject We with its own set of actions and mind.

Under this framework, called Imagined We (IW), collaborators imagine a We agent with a joint mind and "bird's eye perspective" on collaboration, where all individuals are reasoned about as a whole (Nagel, 1989; Tomasello, 2010). That is, We reasons about actions similar to the individual, and would aim to direct cooperators similar to a central controller, just as an individual coordinates their limbs. A true central controller with complete knowledge of the situation could perfectly coordinate all cooperators. However, in reality there is no central controller and it is unrealistic and inefficient for all agents to share all knowledge. Thus, under IW, each agent instead must individually simulate this We agent, making We imagined. IW contains joint mental states — joint beliefs, joint desires, and a joint intention — given observation of joint actions already made by themselves and other agents. The IW mind is constructed by inferring what We believes, desires, and intends based on what We has done (or, as I will explore in Chapter 4, said).

We modeled a coordinated hunting scenario with multiple predators and prey to demonstrate how shared intentionality in IW can help achieve joint commitment to arbitrary goals (Tang et al., 2020, 2022). In this task, all targets were identical and equally rewarding, so that there was no predetermined target; however, predators could be much more successful when they collectively coordinated to hunt together. IW agents bootstrapped commitment to converge on which prey to hunt. Specifically, modeling occurred in three steps:

1. *Intention sampling:* At each time step, each individual samples which prey they believe

is the target based on their version of IW.

2. *Planning:* Given the sampled target, each individual forms a joint plan of how “We” should pursue that target.

3. *Inference:* Each agent takes their own action according to the planned joint policy and observes actions taken by other agents. Using this, each individual does counterfactual reasoning using BToM to ask how they can explain their own and others’ actions if those actions had actually been rationally generated by the supraindiviudal We. This allows individuals to infer which prey “We” most likely wants.

Using this sequence of sampling, planning, and inference, cooperators were able to quickly bootstrap commitment to the same arbitrary intention without even the need for explicit communication. In this setting, IW successfully captures humans’ robust commitment in cooperation: resisting alternative targets, achieving greater quality of hunt, and maintaining a relatively high goal consistency among hunters. This indicates that there is indeed an advantage to shared agency cooperation in this task, and manifesting that advantage requires joint commitment as a stronger constraint for the team behavior.

Approaching this problem from a different angle, recent work has also made a direct comparison to cooperative hunting using a reward sharing model which does not incorporate shared intentionality but instead focuses on a Multi-agent Reinforcement Learning (MARL) framework (Zhao et al., 2022). MARL models, a multi-agent extension to mainstream AI Reinforcement Learning (RL) approaches, traditionally use trial-and-error to approximate the optimal action policy achieved while maximizing long-term expected rewards (Sutton & Barto, 2018). Although this method has been successfully applied to various challenging group coordination scenarios, such as autonomous-driving coordination (Shalev-Shwartz, Shammah, & Shashua, 2016), as well as teaming in Dota 2 (Berner et al., 2019), and StarCraft (Vinyals et al., 2019), it does not consider shared agency and fails to consider the qualitative and semantic differences between cooperation and competition. MARL offers a generic solution to different cooperative (Sayin, Zhang, Leslie, Basar, & Ozdaglar, 2021; Zhang, Yang, Liu, Zhang, & Basar, 2018), competitive (Silver et al., 2017; Sayin et al., 2021; Xie, Chen, Wang, & Yang, 2020; Zhang, Kakade, Basar, & Yang, 2020), or mixed interest settings (Lagoudakis & Parr, 2002) simply by changing the reward function. We argue that to efficiently cooperate or compete, agents need to go beyond the reward structure of the game and consider each other's mental state and possible actions. Reward-sharing on its own only imposes weak constraints on teaming that are insufficient to capture humans' cooperative behaviors.

Zhao et al. (2022) make this case that a shared reward function is not enough for successful coordinated hunting in two parts. First, they show a shared reward is not necessary: selfish predators who only care about their own individual benefits are seemingly able to coordinate with each other to capture a single prey. Moreover sharing a reward is not sufficient: predators that share rewards actually suffer from the free-rider problem. When even a small action cost is added, coordinated hunting breaks down because predators are able to obtain the same reward whether or not they expend effort to contribute to the hunting when that reward is shared. In a hybrid team simulation experiment with multiple prey, the IW model has also been shown to better mimic the intentions of human hunters compared to reward sharing (Tang et al., 2022). Together, these studies offer insights on the potential of using a shared intentionality framework approach to modeling human cooperation.

2.3 Extending BToM for Communication

The second direction BToM has been extended is toward scenarios involving communication, another type of social interaction with a wide variety of situational uncertainties that requires a rich, dynamic exchange of actions and minds. Incorporating communication with an agency-based perspective allows agents to more intelligently solve these situational uncertainties. Paradoxically, adding communication simultaneously introduces its own uncertainty, due to the sparse, overloaded, and flexible manner in which people may communicate. As a result, in order to achieve the immense flexibility inherent in human signaling, cognitively-inspired models of communication must be driven by other strong constraints, many of which lie outside signals themselves, but can be formalized through the structure provided by ToM reasoning. The key insight to extending BToM to capture communication is to treat signaling as a type of rational action which can be planned and reasoned over instead of a code with a fixed mapping. As a result, interpreting signals follows the same framework as generating actions from underlying mental states.

This insight has recently been formalized using the rational speech Rational Speech Act (RSA) (Frank & Goodman, 2012; Goodman & Frank, 2016). Treating signals as a type of rational action means that speakers should be truthful, concise, relevant, and straightforward and a listener should also expect this and interpret the signal in accordance with these properties (Grice, 1975). Under RSA, a signal is used to convey information about beliefs, states of the world, or some referent in a maximally efficient way. In its initial formulation, RSA was used to solve referential signaling games (Lewis, 1969; Wittgenstein, 1953), where a speaker sends a signal with the aim of getting the listener to correctly identify an intended referent among a set of potential referents by reasoning about the linguistic context. Figure 2.2 provides a simple referential signaling game used as a running example.

A rational signal is defined in terms of its utility function: a signal's utility is determined by how it is expected to change the listener's beliefs to reflect an intended referent. This definition provides a mechanism to evaluate which signals are good. Inferring the listener's beliefs in this manner closely mirrors the rational inverse planning that BToM agents perform when reasoning about the underlying mental states that produce observed actions.

Formally, under RSA, a speaker who is pragmatic (p_{sp}), chooses a signal ($signal$) to describe a state of the world $state$, (e.g. the four of clubs). This signal may have multiple referents or interpretations; however, communication is assumed to be produced through a rational decision making process. Parallel to Equation 4.10, a noisy utility maximization

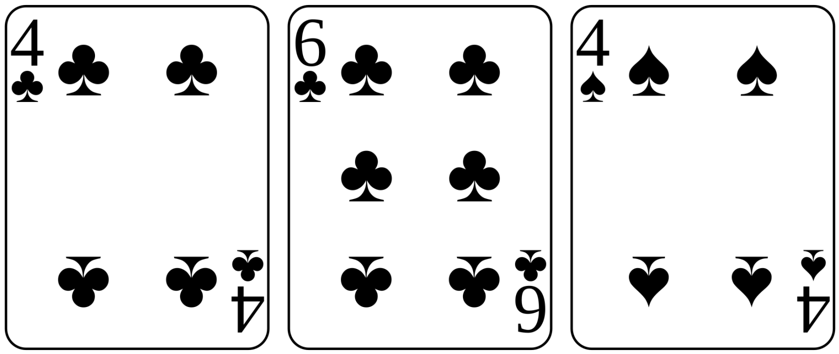


Figure 2.2: Referential Signaling Game Example: Given the cards, the speaker says “clubs.” Clubs can literally refer to two cards – the four and six. However, a pragmatic interpretation would choose the four of clubs using the reasoning – if the speaker had been referring to the six of clubs, “six” would have been a better signal to send.

(soft-max) with $\beta \in [0, \infty)$ representing the degree of rationality, determines the signal sent:

$$P_{sp}(signal|state) \propto e^{\beta \mathbb{E}[U(signal,state)]} \tag{2.4}$$

In order to calculate the utility of a signal, the speaker reasons about how the pragmatic listener (p_{lp}) interprets that signal in terms of the states it could refer to:

$$\begin{aligned} \mathbb{E}[U(signal, state)] &= p_{lp}(state|signal) \\ &\propto p_{sl}(signal|state)p(state) \end{aligned} \tag{2.5}$$

The utility of the signal is defined by how well it is expected to convey a belief about a state to the pragmatic listener. This is estimated by modeling the pragmatic listener's reasoning process: a Bayesian inference about the observed signal. Here, the likelihood function is a simpler, literal speaker (p_{sl}) model which does not have a partner model and simply indicates whether a signal is literally true for a target referent state. This is weighted by a prior over states.

When communicators interact, they can recursively reason about each other at increasingly deep levels; for example, a level-1 speaker models a listener who reasons about a literal (level-0) speaker, but a level-2 speaker considers the listener who models a level-1 speaker. Thus, regardless of recursion depth, the literal speaker acts as an entering point for more sophisticated social recursion. RSA has been incredibly successful in linguistic domains, capturing phenomena including metaphor (Kao, Bergen, & Goodman, 2014), redundancy (Degen, Hawkins, Graf, Kreiss, & Goodman, 2020), and convention formation (R. X. Hawkins, Frank, & Goodman, 2017). However, it has been viewed almost entirely as a language model, focusing on what else others could say. BToM can serve to develop this viewpoint to consider communication in the context of tasks which rely on acting in the environment.

BToM has been applied to model communication in several notable examples, including through communicative demonstrations which illustrates the difference between doing something and showing someone else how to do it. An interesting challenge in these tasks is that the method to communicate completely overlaps with typical actions: agents can take actions in order to achieve their own goals but can also use their actions to communicate information about the environment or task to teach an observer. In this model of pedagogical demonstration (Ho, Littman, MacGlashan, Cushman, & Austerweil, 2016; Ho, Cushman, Littman, & Austerweil, 2021), a demonstrator simply doing her task plans a high utility route to achieve her personal goal given knowledge of the environment. Using standard BToM inverse planning, an observer could then reason about what that goal and knowledge of the environment are. In order to show her goal, the communicative demonstrator incorporates the mind of this observer into her calculation of utility. Traditionally, partially observable MDPs (POMDPs) offer a way for agents to plan over uncertainty in their own beliefs about the environment (Kaelbling, Littman, & Cassandra, 1998); however, here they are being repurposed to plan over uncertainty in the observer's beliefs. This reflects a crossing of minds through recursive reasoning, similar to how RSA speakers make increasingly sophisticated inferences about listeners; the difference here is that the action space is the signal space and thus action utilities are composed of both their effectiveness at achieving goals and their ability to communicate.

In addition, there have been recent efforts to integrate RSA models of communication with actions using BToM. This would allow a separate channel for communication outside of the agent's action space. The major benefit of this perspective is to allow a broader definition of context in communication that includes abilities and actions of agents. Thus, BToM agents' formulation of rational actions in the physical environment allows consideration of non-linguistic context to solve linguistic ambiguity in communication.

One such work examines speaker strategy comparisons that trade-off being maximally informative with maximizing task performance (Sumers, Hawkins, Ho, & Griffiths, 2021). This task combines traditional signaling games introduced with RSA (Lewis, 1969) with the classic reinforcement learning (RL) paradigm multi-armed bandits, which add a more complex reward component (Sutton & Barto, 2018, Chapter 2). A speaker who knows the entire reward structure communicates about rewards associated with features of items, where the sum of all feature rewards is the item's reward. Based on the observed signal, the listener selects an item. Returning to the card task Figure 2.2, the signaler would now be provided information about the payoff of different suits and numbers (e.g. $\spadesuit : +2$, $\clubsuit : +1$, $\diamondsuit : -1$, $\heartsuit : -2$), and could send the listener information about a feature (e.g. $\clubsuit : +1$ or $\clubsuit : +2$ which is a false statement).

This work contrasts three ToM-based frameworks which isolate mental components — beliefs, actions, and the two combined — which allows a fine-grained analysis of the benefits and drawbacks of different signaling strategies. Belief-oriented speakers are RSA signalers, equivalent to Equation 4.4 where *state* now contains all of the speaker's knowledge about rewards and the signaling space is the set of all feature reward pairs. These speakers choose highly truthful but potentially irrelevant signals. In contrast, the action-oriented speaker's utility comes from optimizing how likely the listener is to select the item with the best pay-off from the set of current options based on a model of the listener's actions. Action-oriented speakers choose highly relevant, context-specific signals but often lie to distort beliefs and thus would not generalize well when the current options change. Finally, combined speakers optimize over beliefs that maximize expected rewards by taking into account the rewards of all actions. Combined speakers choose signaling strategies that maximize overall reward in the task, tending to inform (and distort) both high reward and high cost options.

In parallel, we have developed the integration of RSA and BToM — combining beliefs and actions for a formulation of utility similar to the combined speaker to address a completely different phenomenon — to formalize a broader notion of relevance in a helping task. This work has explored how ToM can be leveraged to achieve understanding in a surprisingly sparse but powerful communication modality: pointing (Jiang et al., 2022). Pointing is an interesting communication domain because it is incredibly sparse, indirect, and overloaded, relying entirely on non-linguistic context. Despite this, empirical studies reveal that even prelinguistic infants can flexibly use context to resolve ambiguous social intentions when interpreting pointing gestures (Liebal, Behne, Carpenter, & Tomasello, 2009). A pointing gesture has many potential interpretations, but relevance requires a mutual understanding that what's being communicated should be as relevant as possible (Sperber & Wilson, 1986); thus, the challenge then becomes how to formalize relevancy. In this helping task, an agent who cannot fully observe the environment forms beliefs about their current state after taking each action. Framing this problem as a POMDP allows us to formalize the desirability of taking an action given a belief as a Q-value. The signaler's calculation of relevance then involves a crossing of minds between the signaler and receiver. First, the signaler makes predictions about the receiver's actions using a model of the receiver's beliefs. The signaler assumes that the receiver takes the best possible action given what they know. Second, using their own beliefs, the signaler evaluates that predicted action. Relevance measures the signaler's estimation of how much a listener could *improve* their utility by knowing the information that the signaler has. This formulation reflects paternalistic helping, a phenomenon where a person who knows better rejects someone's request "for their own good," even though the requester may not feel this way. Modeling work in Chapter 3 of this dissertation also takes this approach to capture the formation of instantaneous communicative conventions using shared context.

2.4 Conclusions: Shared Agency Cooperative Communication

Thus far, we have reviewed how individual BToM can be extended for shared agency to model cooperative cases and extended by treating signals as rational actions to model communication. Integrating these two lines of research has a strong theoretical motivation that stems from psychology and evolutionary biology. One influential theory on the origin of communication is that it arose as an adaptation to support complex underlying social cognition in humans, specifically in support of increasingly complex cooperation (Tomasello, 2010). As a result, communication is simply another type of social tool allowing people to get things done together (Bruner, 1985; Vygotsky, 1978). Thus, communication should be viewed in terms of its use, specifically the ways it can be cooperatively used to align perspectives. These two parallel lines of research are unified in a shared agency model of communication capable of handling uncertainty in signals in Chapter 4. This will broaden the scope of considered communication tasks from helping or observing to include partners who are equally able to act and make decisions in the environment.

CHAPTER 3

Overloaded Communication as Paternalistic Helping

Even simple, ambiguous signals can have rich interpretations when viewed as part of an interaction in a shared environment. We create a model called Paternalistic Communication (PaCo), designed to formalize aspects of this context by combining an existing modeling framework for overloaded signaling — Rational Speech Acts (RSA) — with an agentic Theory of Mind (ToM) model. This integration allows signals to be processed in conjunction with common ground in a principled manner using task-dependent action utilities. This modeling perspective treats communication as a way to coordinate diverging perspectives in a cooperative setting. Under the PaCo framework, a speaker decides what to say by predicting their partner's response based on public information in the common ground and then evaluates those responses using private information in their own mind. We demonstrate the potential of this framework using an existing case study with context-based, ambiguous signaling. Through a set of simulations where we compare PaCo to RSA, we replicate the flexibility of human performance and extend beyond the original task to show how common ground constraints and additional levels of modeling recursion affect performance.

3.1 Background

You're walking with a friend on a freezing winter night when your friend yells "careful!" You look down and observe a patch of black ice underfoot. Without context, "careful!" can mean countless things; however, in context, this single word sets off a rich inferential process: What is your friend referring to? (the ice) How should this knowledge change your beliefs? (the ground will be slippery) How should this change your actions? (falling hurts, so tread carefully). One traditional approach to modeling communication assumes that words and their meaning have a one-to-one mapping, predefined outside of the current exchange (Shannon, 1948; Valiant, 1984) which would fail at explaining this example. “Careful!” cannot be mapped from a codebook, as the meaning of the signal is derived directly from the context of the current situation. At another time, I could easily be referring to broken glass from a dropped jar or a large pothole in the road. However, instead of an encoding and decoding process, human communication is highly dependent on understanding what is relevant in the current context (Sperber & Wilson, 1986), allowing us to be incredibly successful at expressing rich meaning using sparse, overloaded signals. In this work, we propose a model of signaling that targets how the context of the situation can help solve signal ambiguity.

Sending a sparse signal is often enough to impact a listener’s mind in sophisticated ways. This is because a speaker counts on a listener to be intelligently using her own mind — full of beliefs, desires and intentions — to interpret a signal’s meaning in context of mutually known information: the common ground. In turn, the listener trusts that the speaker has chosen a signal by first considering how an intelligent listener might act in response. The examples above highlight that communication is often used as a cooperative tool for helping (Tomasello, 2010). However, communication is a unique type of helping for two reasons. First, it is not the same as instrumental helping because instead of taking actions that change the world, communicators send signals to change the mind. Second, communication requires a coordination of minds. Communicators simultaneously track what is shared and what is private (Heller, Parisien, & Stevenson, 2016), which requires agents to coordinate divergent minds. In order to coordinate these minds in a cooperative setting, we turn to a type of helping that has already been studied: paternalistic helping. In the following sections, we introduce a set of components that allows us to build a flexible model of communication that can be integrated using the principle of paternalistic helping.

Overloaded communication in a visual scene has also been studied through empirical research in psychology (Grosse, Moll, & Tomasello, 2010; Tomasello & Haberl, 2003). In one notable experiment which we use as a modeling case study, cooperators are demonstrated to use and understand overloaded signals to form instantaneous conventions that change flexibly, depending on context (Misyak et al., 2016). In this task to collect rewards (bananas) and avoid punishments (scorpions) hidden in boxes, one partner knows where rewards are located, but cannot reach them while the other does not know the boxes' contents but may use axes to open them and collect whatever is inside. The knowledgeable speaker can signal by placing tokens on boxes, which can equivalently be interpreted as "open" or "avoid" that box. Across trials, shared knowledge and available actions are manipulated. Using this, partners who have never interacted with each other before are able to successfully and flexibly use and interpret the meaning of the tokens. In this work, we build a computational model to capture the flexibility seen in this banana example by treating communication as a type of paternalistic helping.

3.1.1 Common Ground

Common ground is mutually shared, public, and transparent knowledge that can be assumed between communicators. Keeping track of this information is critical to communication (Clark, 1992). Initially, this was proposed as a recursive statement: information in the common ground must only contain things I know you know that I know ... *ad infinitum* (Lewis, 1969; Schiffer, 1972). However, the feasibility of establishing common ground through infinite recursion has been challenged by the speed and sparsity of everyday exchanges (Sperber & Wilson, 1986) as well as the fact that speakers use self-repair to add information when they detect their partner is not on the same page (W. J. Levelt, 1983). Heuristics have since been proposed to bypass much of the recursion (Clark & Marshall, 1981), allowing communicators to assume some common knowledge. Here, we treat communication as a means to add information into the common ground. In turn, the existing common ground also acts as a means to narrow the interpretation of a signal (Clark & Brennan, 1991; Clark & Marshall, 1981) which allows people to communicate more efficiently.

When knowledge is already in the common ground, there is no longer a need to talk about it, reducing the length of communication. Returning to the ice example, because speakers assume shared knowledge and only need to talk about things that are new, "there might be ice underfoot which could cause you to slip and fall which would hurt so be careful" achieves the same effect as a simple "careful." Similarly, in the banana collecting task, the total number of rewards is sometimes in the common ground so tokens are placed already assuming that knowledge. In addition, a new signal should be interpreted in the context of the already established common ground, making its meaning flexible. Thus, changes in the common ground allow the same token placed on a box with a banana in one situation and a scorpion in another.

Evidence from developmental studies reveal that even infants are able to use common ground to resolve ambiguity in communication both when interpreting and generating an ambiguous signal. Infants as young as 14 and 15 months old use past shared experiences to resolve ambiguous requests for an object (Ganea & Saylor, 2007; Moll, Richter, Carpenter, & Tomasello, 2008). Moreover, infants interpret pointing to a fixed object differently, resolving ambiguous social intentions of the pointer based on the previous shared interaction with them (Liebal et al., 2009). Pre-linguistic infants also use common ground to resolve ambiguity when generating signals, namely through their use of pointing. One-year-old infants point more for adults who are ignorant, indicating that these infants have an understanding of what is and isn't in the common ground (Liszkowski, Carpenter, & Tomasello, 2007). Pre-linguistic infants even use pointing to note the absence of items (Liszkowski et al., 2007, 2009). Pointing to something out of place first requires an expectation of where things normally are in the common ground. The violation of this expectation, not the current shared experience, is the referent of this gesture. Collectively, these studies demonstrate that even early on, humans have the capacity to reason richly about incredibly sparse communication, which go far beyond simply inferring what the referent is.

These effects are also seen in adults where coordinating common ground occurs on many different scales and from many different sources. Broadly, individuals can share societal cultural norms and shared domain knowledge (Isaacs & Clark, 1987). At a fine-grained level, individuals may share a history of past experiences and interactions (Fussell & Krauss, 1992), or even just an immediate shared visual context (Richardson, Dale, & Kirkham, 2007; Hanna, Tanenhaus, & Trueswell, 2003). Common ground even allows individuals to coordinate on a social level without explicit communication, relying instead on "virtual bargaining" and salience (Misyak, Melkonyan, Zeitoun, & Chater, 2014; Schelling, 1960).

Because it can come from the environment itself, common ground allows overloaded communication to be instantaneously interpretable without requiring a past history of interaction (Clark, 1996; Tomasello, 2010). In turn, this allows it to serve as a natural anchoring point for future communicative exchange. Here we focus on this type of instantaneously formed common ground which relies on an intuitive understanding of others' minds which is largely universal and developed early on (Wellman, 1992). Of course, common ground is not static, but instead interacts with communicative exchange. This process is a dynamic feedback loop between signal and common ground: each communicative contribution is context-shaped and context-renewing (Heritage, 1984).

3.1.2 Flexible Linguistic Pragmatics

The literal meaning of an utterance is not always enough to resolve a signal. Instead, pragmatics — taking into consideration the context of the current exchange — is often the key to understanding what someone means (Levinson, 1983). Previously, we've discussed why common ground gives context to a situation. Here we briefly review a framework called the Rational Speech Act (RSA) model which has seen much recent success as a formalization of linguistic pragmatics (Frank & Goodman, 2012; Goodman & Frank, 2016). Specifically, we introduce an extension which allows for different speaker types who have varied communicative goals. In the context of this work, we simply refer to this variant as RSA. In the banana example, this can capture the opposite usage of tokens to communicate when to open or avoid a box; more broadly, this extension has also been used to describe linguistic phenomena such as hyperbole (Goodman & Frank, 2016; Kao et al., 2014).

Under RSA, a pragmatic speaker ps chooses a *signal* from a set of possible signals and signaling type c to describe a target referent or world state w. Signals are treated as a type of rational action, subject to a utility soft-maximization, where $\beta \in [0, \infty)$ represents the degree of rationality. Here, the utility of a signal can be calculated by reasoning how a pragmatic listener pl will interpret that signal:

$$P_{ps}(signal, c|w) \propto e^{\beta P_{pl}(w,c|signal)} \tag{3.1}$$

The pragmatic listener models signal interpretation using Bayesian inference, which requires a likelihood function, here a simple, generative literal speaker model ls. Finally, the prior term is defined over both speaker type and state, which are assumed to be independent in this task context.

$$P_{pl}(w, c|signal) \propto P_{ls}(signal|w, c)P(w)P(c) \tag{3.2}$$

In order to generate this interpretation, the listener also needs a speaker model in the form of a likelihood function. To start with a simple entering point that doesn't require further reasoning, the speaker is modeled as literal ls, uniformly sending true signals according to a lexicon ($Lex(.)$). The lexicon is an indicator truth function of whether a signal is consistent with the referent state w given the speaker type c, returning one when consistent and zero otherwise:

$$p_{ls}(signal|state) \propto e^{\beta Lex(signal,state,c)} \tag{3.3}$$

An important property of RSA is its social recursive reasoning. In this example, a speaker models a listener modeling a basic speaker. However, it is straightforward to add additional layers of reasoning about your partner, building up a "cognitive hierarchy" (Camerer, Ho, & Chong, 2004).

RSA is grounded in linguistic theory, including Gricean cooperative logic (Grice, 1975). Grice's insight is to treat conversation as cooperative: interlocutors should speak in ways that are maximally efficient and rational to make oneself interpretable, which serves as a constraint to how a signal can be understood. As a result, RSA has successfully modeled a variety of linguistic phenomena in communication, such as implicature (Goodman & Stuhlmüller, 2013), vagueness (Lassiter & Goodman, 2017), and convention formation (R. X. Hawkins et al., 2017). While some recent work has begun to develop in the direction of adding action context (Sumers et al., 2021) or grounding pragmatic signals within a utility-driven task (McCarthy, Hawkins, Wang, Holdaway, & Fan, 2021), RSA has primarily been used in purely linguistic settings. In these cases, speakers have the communicative goal of describing a referent by reasoning about how different signals are expected change the listener's beliefs.

As a generic model of communication, RSA requires an informative means to evaluate a signal. Conceptually, in its current formulation, RSA provides the required structure by tying communication to concrete utilities: in the black ice example "careful" can mean different things and we should choose the meaning that allows for the best utility. Similarly, in the banana example, we can interpret the token by reasoning about whether interpreting it as "open" or "avoid" can achieve a higher utility. RSA's approach to deriving the utility of a signal is based on its ability to efficiently convey the features of its target referent state, $p_{pl}(state|signal)$, to the listener. However, describing features of a target is only one part of human communication.

In reality communication is often several steps more indirect than this; we communicate about not only referent states (What?) but also social motivations (Why?) and interactions in the shared physical environment that can achieve those motivations (How?). In the black ice example, “careful” literally pertains to an agent’s actions and is motivated by helping a friend avoid a large negative action utility (slipping on ice). In the banana example, one condition manipulates the receiver’s abilities which does not change the state of the world — where the bananas are — but still affects how people send and interpret signals. To consider more generic settings where communication is indirect, we need to integrate the utility of a signal with an intelligent model of agency. This will allow us to consider communication in terms of what agents can do, what agents want to do, and how they can act to achieve it.

3.1.3 Bayesian Theory of Mind

To provide the necessary cognitive infrastructure for a model of agency, we turn to Theory of Mind (ToM) which has been extensively studied in cognitive science (Premack & Woodruff, 1978; Wellman, 1992). Particularly, we focus on the Bayesian formulation of ToM for modeling purposes. ToM provides a natural mechanism for rational action planning and inverse action interpretation. When deciding how to act, one should rationally take actions that achieve desirable utilities with respect to their underlying mind, which can be broken down into different mental states including beliefs, desires, and intentions:

$$P(action|mind) \propto e^{\beta \mathbb{E}[U(action, mind)]} \tag{3.4}$$

In addition to action planning, ToM is a model of action interpretation. Bayesian inference allows an observer to reason over which underlying mental states an agent is most likely to have given their observed actions (see Equation 3.5). Bayesian ToM models of agency have been widely and successfully adopted in the last few decades to infer physical goals (Baker et al., 2009; Jara-Ettinger, Gweon, Schulz, & Tenenbaum, 2016), social goals (Ullman et al., 2009), and joint beliefs and desires (Baker et al., 2017) from observed actions. In addition, ToM also has great potential in cooperation, using joint planning (Kleiman-Weiner et al., 2016; Shum et al., 2019).

$$P(mind|action) \propto P(action|mind)P(mind) \tag{3.5}$$

ToM agents act rationally according to the expected utility of taking different actions, but where do these action utilities come from and why do they matter? In the banana example, there are no explicit action costs, but opening boxes with bananas leads to positive rewards while scorpions lead to negative ones. Even preschoolers can understand basic costs of acting and rewards associated with different preferences and desires using a Naive Utility Calculus (Jara-Ettinger et al., 2016; Liu, Ullman, Tenenbaum, & Spelke, 2017). These cost-reward comparisons constrain the scope of affordable actions, allowing cooperators to resolve ambiguity (Jara-Ettinger, Floyd, Huey, Tenenbaum, & Schulz, 2020). By using ToM reasoning to understand communication, we have a way to evaluate signals by connecting them back to a task with well defined utilities.

3.1.4 Utility of a Signal

In addition to action utilities, communicators also need to evaluate the utility of a signal. However, the utility of a signal is not as straightforward to define as it may initially seem. Unlike instrumental actions, signals do not directly change the world. Instead signals are designed to change our mind — affecting beliefs, desires, and even actions. If signals cannot change the world, how can we define their utility? We we start with the insight that changing one's mind can lead to different rational behavior. Given the changed mind, we can make a prediction about what a rational agent will do. These actions serve as the ultimate consequence of sending a signal. Thus, by grounding our communicative interactions in action consequences, we can measure the utility of sending a signal. By considering the entire mind, communication becomes more than reasoning about beliefs, instead, it extends to what agents want and will do given an evaluation of the consequences of those actions. As a result, this agency-based model can actually significantly enrich the notion of a signal's utility.

The focus on the tight connection between communication and actions, showing how communication can be used to do things, has long reaching roots in philosophy, in language games (Wittgenstein, 1953). Austin (1962) developed this further through “speech acts,” where signals can be viewed as a type of rational action that can do things. This same approach has been adopted in Artificial Intelligence (AI) as a way to derive the utility of a signal by grounding it in the resulting action. In these approaches, the value of sending a signal is equivalent to how good the actions taken as a result of hearing it are, compared to the action that would have been taken without the signal.

One such modeling framework is Recursive Mind Modeling (RMM) (P. J. Gmytrasiewicz & Durfee, 2001). The key insight of RMMs is to derive the value of communication by grounding interaction in a task in the environment. Under an RMM framework, a signal is designed to change the decision-making situation agents are in. Namely, a signal can reduce an agent’s belief uncertainty about the reward structure of the environment, allowing them to take actions that are more likely to lead to beneficial outcomes. RMMs are promising in formalizing signal utility but bypass a critical challenge of human-like communication: how a signal can be mapped onto its meaning. RMMs generally treat the mapping between signal and meaning as one-to-one and fixed: the uncertainty in the expected utility of a signal comes only from noise in the communication channel. Treating uncertainty only as a probability of transmission error ignores the additional — and much more challenging — uncertainty that comes from trying to understand an ambiguous signal’s meaning. Because the mapping between utterance and meaning is fixed, there is no need to reason about what that signal should be, because it can only mean one thing. To move beyond codebook mapping communication, we must integrate the RSA framework, which can handle this type of overloaded and flexible mapping, with the task oriented definition of signal utility seen in AI works.

3.2 A Paternalistic Perspective on Communication

When communicating, a final issue that arises is how to coordinate the communicators' diverging minds, which can contain asymmetric information. To understand this, it is useful to start by thinking of cooperative communication as a type of helping. When sending a signal, speakers help listeners by providing new, relevant information not already in the common ground. The utility calculation underlying this signal requires both prediction and evaluation of a listener's actions. There is already an elegant solution to the perspective coordination problem: Smithian sympathy. Smith demonstrated that sympathy encompasses more than resonating with others' feelings (Smith, 2010 [1759]). One can feel bad for a dead person or sad for someone who has dementia but feels happy.

More recently, this idea has been extended to paternalistic helping where a person who "knows better" overrides another individual's actions or requests "for their own good," even though the requester may not feel this way (Jacobsson, Johannesson, & Borgquist, 2007; Sibicky, Schroeder, & Dovidio, 1995). Paternalistic helping develops early: children as young as five years old are able to override another child's desires by considering the negative consequences of that desire (Martin, Lin, & Olson, 2016). The key to this paternalistic perspective is that a person can predict others' actions by ToM reasoning using common ground knowledge, but that person will *evaluate* predicted actions according to their private beliefs. Taking a paternalistic perspective has previously been successful for modeling when helpful pointing is relevant in a classic AI task (Jiang et al., 2021). Similarly, we adopt this approach as a way to coordinate the minds of speaker and listener.

3.2.1 Paternalistic Communication (PaCo) Modeling Framework

Paternalistic helping acts as a binding agent between common ground, RSA, ToM, and signal utilities derived from actions to process flexible, sparse communication in a holistic manner. For this reason, we call our modeling framework Paternalistic Communication (PaCo). A pragmatic paternalistic signaler chooses what to say by evaluating the utility of different signals, equivalent to the pragmatic RSA speaker (Equation 3.1) with a generalized utility function. This generalized value of a signal for an intentional agent is linked back to action through a task in the environment. The speaker creates an expectation of how good a signal will be by predicting how a receiver, upon hearing the signal, will act. Then the speaker evaluates how good that action is compared to other actions generated by other potential signals:

$$\mathbb{E}[U(signal, mind)] = \mathbb{E}_{P(a|signal,c)}[U(a, mind)] \tag{3.6}$$

There are two terms needed to connect signals to actions. First, we measure how signals change the mind. This quantity, $P(mind_{cg}|signal, c)$ can be derived from inverse planning in ToM where signals are treated as a type of rational action (Equation 3.5) and is similar to modeling a RSA listener (Equation 3.2), but is capable of reasoning more generically over other components of the mind. Second, we measure the rational action outcomes that arise from different mental states. Here, $P(a|mind_{cg})$ can be derived from ToM rational action planning (Equation 3.4).

$$P(a|signal, c) = \sum_{mind_{cg}} P(mind_{cg}|signal, c)P(a|mind_{cg}) \tag{3.7}$$

Here we assume that the signal and speaker type are independent from other components of the mind. The integration of common ground, ToM and RSA under the paradigm of paternalistic helping gives a flexible, context driven approach to overloaded communication.

These two components, action prediction and signal evaluation, reflect the different perspectives of paternalistic sympathy. Action prediction is done with respect to common ground information while action utilities are evaluated according to the speaker's privileged information about the world. Thus, a signaler predicts what actions a receiver is likely to take by simulating the *receiver's* mind, which constrains which actions are rational to take downstream. The underlying receiver mind is denoted $mind_{cg}$ to emphasize the role of common ground information that can create strong interpretation constraints before communication even occurs.

The integration of these modeling components gives a flexible, context driven approach to overloaded communication. First, common ground cues provide the necessary perceptual context to constrain sparse signals. Second, RSA gives a framework for flexible signal production and interpretation through creating a model of one's communicative partner that can predict how signals will be interpreted and evaluate how good those outcomes are. Third, ToM expands the definition of *how* to model a communicative partner, allowing common ground information to be processed differentially and grounding signals into tasks in the environment. Finally, because of the link from signal to action, the utility of a signal can be more naturally expressed in terms of intuitive costs and rewards of acting in the environment.

3.3 Case Study Modeling

Here we demonstrate the potential of PaCo by modeling a case study with impromptu, overloaded signaling from Misyak et al. (2016). Through a non-linguistic cooperative communication task, the authors empirically demonstrate that humans can coordinate to form instantaneous conventions using contextual environmental cues in the common ground, even when a signal can mean opposite things ("go to" or "avoid" a location). We provide a computational account of these behaviors as a special case of ambiguous communication that is captured by PaCo and compare it to a baseline model: the version of RSA introduced previously which allows varied communicative goals. The key difference between these models is that in PaCo context includes both beliefs about world features and how an agent can act based on the knowledge of those features; whereas, RSA is only able to consider beliefs.

3.3.1 Task

During each trial, each pair of participants saw three boxes, each containing either a banana (a reward) or a scorpion (punishment). The goal of this task was to collect bananas while avoiding scorpions. The participant who played the role of signaler knew which boxes contained bananas and scorpions but could not open them. The receiver had no information about the contents of boxes but could use axes to open them. At each trial, the signaler had a fixed number of tokens which could be placed on top of boxes to provide information for the receiver to see (a maximum of one token could be placed on a box). The receiver had a fixed number of axes with which to open boxes. Both individuals knew how many tokens and axes were available, making this common ground information (see Figure 3.1). The total number of bananas and scorpions was either shown in the common ground or occluded by a wall to hide that information (wall not shown in figure). Communication in this task was always fully overloaded because placing a token on a box had two opposite interpretations: "go there" or "avoid that."[1]

There are four key conditions that can be used to highlight how humans flexibly convey meaning across context: Two Token, Inversion, One Ax, and Wall. In all four of these conditions, two of the boxes contain rewards and the third contains a punishment, but the receiver does not necessarily know this. In the Two Token condition, the signaler has two tokens, the receiver has two axes, and the number of bananas and scorpions is known. The Inversion condition is identical to the two token condition, except the signaler is only given one token to convey the location of the two rewards. When, in addition to one token, the receiver also has only one ax, this is the One Ax condition. Finally, the Wall condition occludes the number of boxes containing rewards and punishments from the common ground, while leaving the signaler with one token and the receiver with two axes. Conditions are summarized in Table 3.1.

[1] The true space of meanings could be infinite, containing conventions such as "avoid the box to the left of the token"; however, following the original experiment, we focus on these two which are most direct.

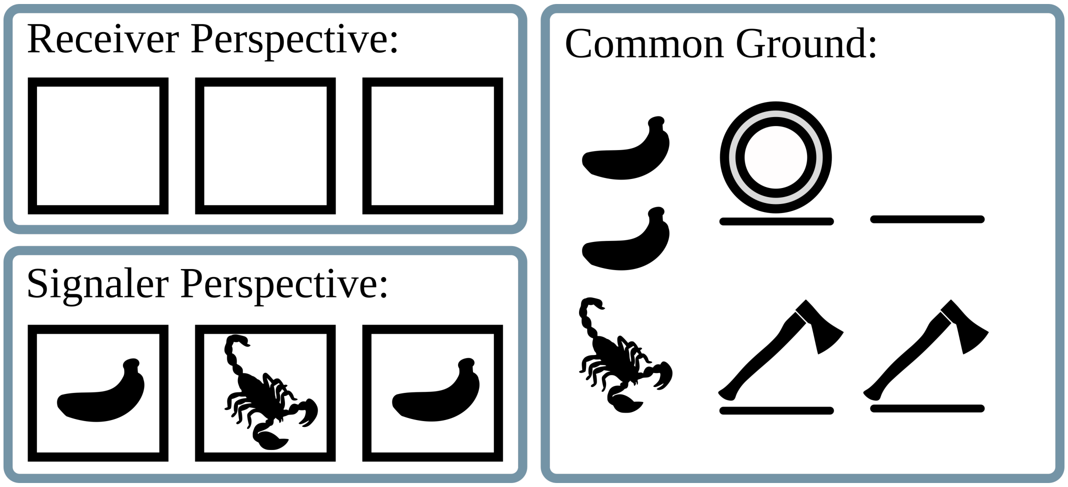


Figure 3.1: Schematic of Inversion condition setup for Misyak et al. (2016). (Left) Agent specific knowledge: the receiver sees the three boxes but not their contents whereas the signaler sees the contents of each box. (Right) Common ground information: both participants see how many tokens are available to the signaler, how many axes are available to the receiver, and the total number of boxes containing bananas or scorpions (which is sometimes hidden behind a wall).

In order to focus on instantaneously formed conventions from impromptu communication, we model a version of this game played without live interaction between signalers and receivers. This limits learning and rapport building between individuals, as the sequence of signals encountered by the receiver comes from multiple, anonymous signalers, collected at an earlier time. In addition, no feedback was given to either player about the outcome of how they chose to produce and interpret a signal's meaning. See Experiment 2 in Misyak et al. (2016) for details.

This experiment emphasizes the importance of common ground as context to solve ambiguous communication. Here, the common ground is composed of shared information about the number of rewards, number of axes, and number of signals in the task. The possible contents of the boxes form the space of beliefs, utility maximization principles of collecting bananas and avoiding scorpions defines the desire, and which boxes the receiver can open

Condition	Tokens	Axes	Wall Present
Two Token	2	2	False
Inversion	1	2	False
One Ax	1	1	False
Wall	1	2	True

Table 3.1: Key experimental conditions in Misyak et al. (2016)

given available axes form the action space.

Under PaCo, these beliefs and desires compose $mind_{cg}$ and operationalize the utility $U(.)$ of an action, while the number of axes the receiver has defines the space of possible actions $a \in A$. The number of tokens defines the space of all possible signals, while the number of axes, rewards, and signals define a prior over the shared content of the common ground. Communication in the task is always fully overloaded because placing a token on a box can have two opposite interpretations: "go there" or "avoid that," depending on the speaker type c. Thus, disambiguation occurs on a trial-by-trial basis as receivers flexibly and jointly infer the tokens' meaning and, as a direct consequence, the boxes' contents. In this context, $mind_{cg}$ is effectively equivalent to RSA's notion of w in Equation 3.1, but can naturally generalize to include uncertainty in joint desires. For example, while we consider the benefit of a banana equivalent to the cost of a scorpion following the original experiment, this could be compared to a model of a speaker who is risk averse and cares more about avoiding scorpions than collecting bananas. The critical difference between these two formulations is that under PaCo, the speaker aims to maximize the receiver's action utility. As a result, the available actions provides a constraint for $mind_{cg}$ which influences which signal a speaker chooses; in contrast, in RSA the speaker aims to directly convey maximal information about w, the state of the world, without considering how a receiver can act on it.

3.4 Simulation 1: Capturing Human-like Use of the Same Signal for Opposite Meanings

In Simulation 1, we aim to capture humanlike flexible signaling behavior across the four key conditions introduced above. In particular, we highlight cases where PaCo and RSA give different predictions. While RSA can use signals flexibly to maximally resolve beliefs about the world state, this may not always be the optimal communication strategy: a fact which humans are sensitive to.

For example, the only difference between the Inverse condition and One Ax condition is that the receiver can open two boxes in the former but only one in the latter. Humans are more likely to use their single token to denote a punishment in the Inverse condition (providing maximal information about the world) and to denote a reward in the One Ax condition (providing an action directive). In this second case, because participants can get at most one reward with only one ax, extra information about the second reward is extraneous. Signaling in opposite ways in these two conditions relies on the signaler's expectation formed through ToM action prediction that the receiver will act differently based on the rational integration of beliefs and available actions. We predict that PaCo will robustly capture human-like flexible use of tokens in these conditions as well as the other key conditions tested in the original study.

3.4.1 Methods

3.4.1.1 Task Specification

To translate the task's goal into an explicit utility calculation, we assign a positive value (+1) for each banana and a negative cost (-1) for a scorpion. Unlike traditional RSA, this cost ratio could naturally vary using PaCo; however, this is not a factor considered in the original behavioral experiment, thus we choose a fixed constant where the benefit of choosing a banana is equivalent to the cost of choosing the scorpion. Also following the original study, there is no explicit cost of using more tokens, if available. PaCo and RSA can both be characterized by two free parameters: β and $P(open)$. β offers an estimation of how rational an agent is; we assume partners are equally rational. $P(open)$ represents the prior distribution over signaler type. We focus on the two types primarily employed by humans: $c \in \{avoid, open\}$. An open-type signaler only places tokens on bananas while an avoid-type signaler only places tokens on scorpions. The prior over beliefs $p(mind_{cg})$ is uniformly split across all possible assignments of bananas and scorpions; when there is no wall, all assignments inconsistent with the common ground beliefs are given zero probability. To test the robustness of the models, we compare model predictions of how the signaler will act under a wide range of parameter combinations (β = [1, 2, ..., 17], $P(open)$ = [.4, .425, ..., .675, .7]). For each combination, we let the two models play the same task as seen by humans in the original experiment.

3.4.1.2 Descriptive Statistics

We define a single scalar to describe the similarity between model and human generation of a signal. An averaged root-mean-squared-error ($\overline{RMSE}$) quantifies how closely the model approximates human signal generation, where a smaller $\overline{RMSE}$ indicates better agreement between human and model. For a particular condition, we first categorize behavior into the two strategies a signaler could employ and take the RMSE between model x and human $x*$ distribution. Then, across the four conditions, these RMSEs are averaged to get obtain $\overline{RMSE}$:

$$\overline{RMSE} = \frac{1}{4} \sum_{m \in Condition} \left(\sqrt{\frac{1}{2} \sum_{x \in open, avoid} (x_m - x_m^*)^2} \right) \tag{3.8}$$

3.4.2 Results

To understand how robust each model is to changes in hyperparameters, we calculate the $\overline{RMSE}$ across the grid of β and meaning priors for each model, summarized in Figure 3.2. To compare overall tolerance to parameter changes between the two models, we conducted a one-sided Wilcoxon signed-rank test for matched-pairs. Under equivalent conditions, the median error under PaCo is significantly smaller than RSA (W = 630, $p < .001$). This supports PaCo's robustness across a wide range of parameters, and suggests that these properties are not the product of over-fitting human data, but rather, a specific example of a general class of phenomena a paternalistic perspective is capable of handling.

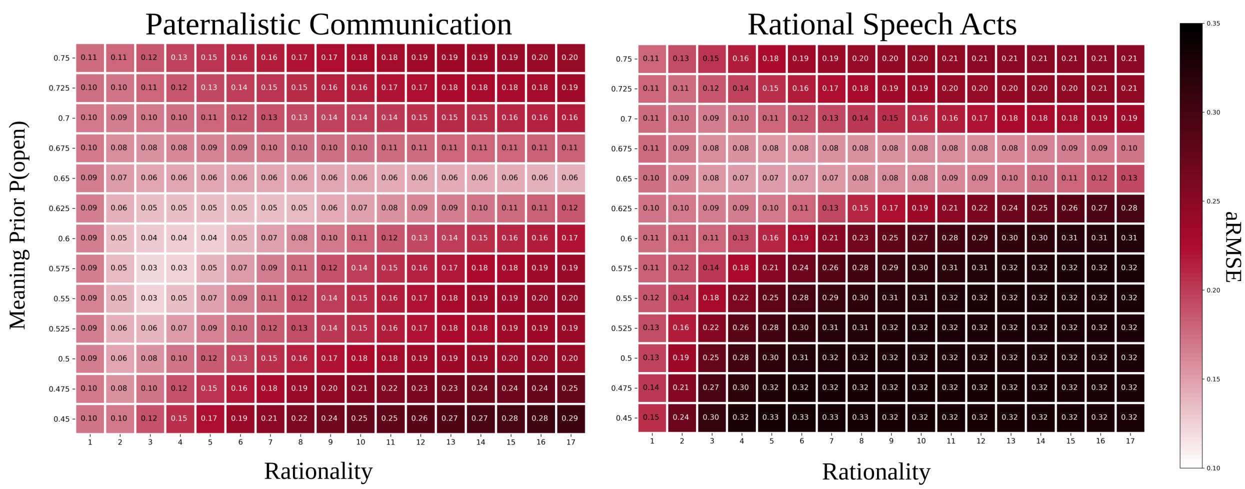


Figure 3.2: PaCo and RSA heatmaps of $\overline{RMSE}$ for key trials: The $\overline{RMSE}$ for each model and parameter combination is represented as a color intensity in the heatmap. Lighter colors represent smaller error or, equivalently, better agreement between human and model. PaCo systematically reduces error better than RSA across parameter combinations.

Beyond the overall fit, we look at strategies employed in the four key conditions, paying specific attention to differences between Inversion and One Ax, where humans tend to change their strategy between conditions. To do this, we select the parameter set that best approximates human strategies in terms of error minimization for each model (PaCo: $\beta =$

3, P(open) = .575 results in $\overline{RMSE} = 2.94 \times 10^{-2}$, RSA: $\beta = 5$, $P(open) = .65$ results in $\overline{RMSE} = 7.05 \times 10^{-2}$).

Like humans, PaCo is sensitive to the common ground: how many signals and axes were available and the presence/absence of the wall, and instantaneously changes which strategy is dominant between the Inversion and One Ax conditions (Human $P(open)$: Inv =.42, One Ax = .63; PaCo: Inv = .47, One Ax =.57). In contrast, RSA fails to make this strategy switch or even distinguish between these conditions (Inversion = One Ax = .58). Full behavioral pattern in key conditions are shown in Figure 3.3.

Using ToM, PaCo can integrate and reason about common ground information that relates to different components of the mind. Here shared information about how many boxes contain bananas puts constraints on possible beliefs about the state of the world prior to communication. Critically, the information about how many axes are available in the environment also relates to the mind as a constraint to the action space. A utility driven choice of a signal meaning under constraints of the signal space processes both of these common ground components jointly. Next, we explore how common ground information produces flexible signaling through simulation results beyond those in the original study.

3.5 Simulation 2: Understanding the Effects of Action Driven Utility

Previously, we demonstrated that PaCo and RSA behave differently at capturing human signaling flexibility; here, we explain these differences by focusing on the contributing contextual factors from a theoretical standpoint. To examine how these models produce different behaviors, we divide context into two separate sources of uncertainty within the common ground. First is the world space, manipulated by the presence or absence of a wall, and second is the action space, manipulated by the number of available axes. Through a set of highlighted scenarios, we examine how these elements contribute differently to achieved

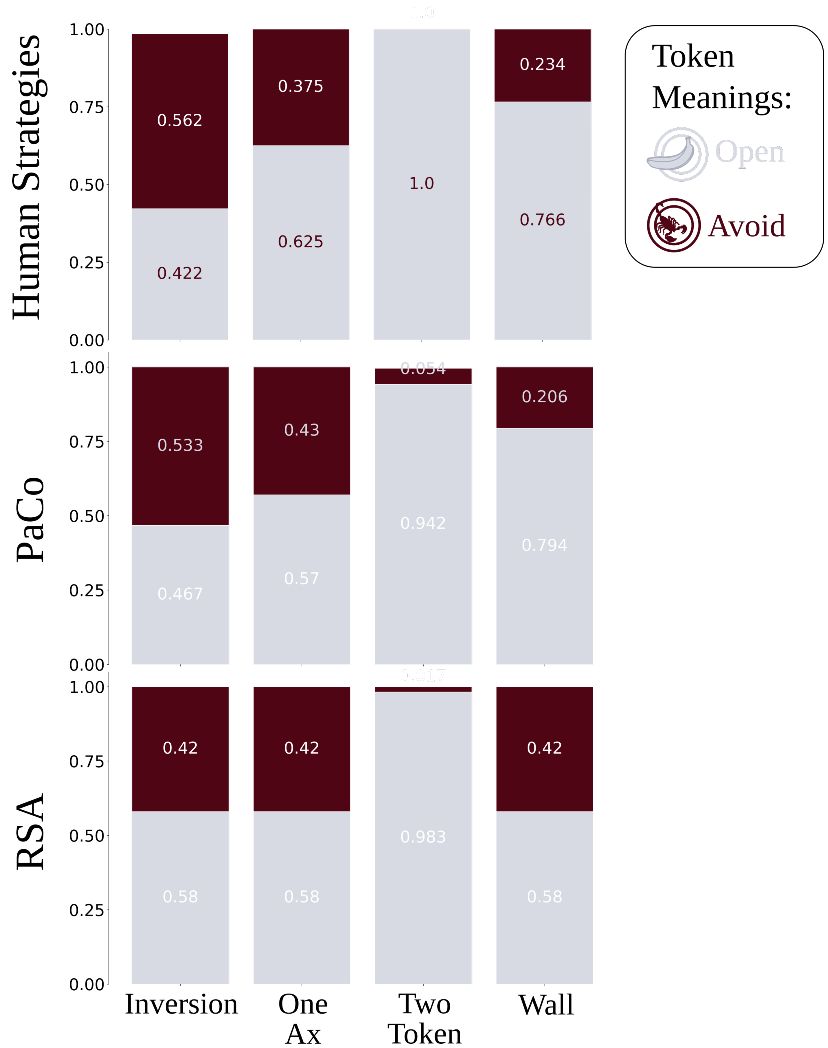


Figure 3.3: Strategy breakdown comparison between humans, PaCo, and RSA under optimal parameters. Trials where the predicted strategy is to place tokens on bananas in grey and on scorpions in navy.

utility in the task as well as signaling behavior.

The major difference between these models is that PaCo derives its utility from how desirable action outcomes under the task are expected to be whereas RSA focuses on minimizing the uncertainty in a listener's beliefs. Thus, we expect RSA to be sensitive only to the world space knowledge, not restrictions to actions. On the other hand, how well PaCo performs should depend on whether considering the receiver's action space can act as a constraint on signaling. When the receiver is limited to few actions, these limitations propagate to what the signaler should say. A signaler should not tell the receiver to do something known to be impossible. However, when the receiver is capable, with a large action space, jointly considering actions and the world can increase uncertainty, making it more difficult to understand what the signaler is trying to communicate. Thus, we expect PaCo's action-based reasoning to become more important in cases where the world state is highly uncertain.

3.5.1 Methods

3.5.1.1 Task Specification

We use the same task utility structure as before, adding a small cost (-.1) per token used to encourage shorter signals. In addition, to look at how performance varies across scaled-up environments, we expand the world to have five boxes. Token meaning priors are set at the optimal ones that match human performance in Simulation 1, and the models are set to high rationality ($\beta = 20$) to emphasize theoretical performance. The number of axes are manipulated (1, 2, 3, 4), with and without a wall. The number of tokens are set to be high (3, 4) which ensures that a signaler has the means to send a longer signal if desired. Similarly, the number of rewards are set to be high (3, 4) which ensures the possibility of achieving a high utility. We sample N=250 environments for each combination of wall and number of axes.

3.5.1.2 Descriptive Statistics

To test how PaCo and RSA communicated using different strategies, we used the Kullback-Leibler (KL) divergence between P, the true belief that the signaler privately knows (i.e. which boxes contain scorpions as opposed to bananas) and Q, the receiver's belief posterior about the box contents to describe the uncertainty over the set of possible beliefs M. Because the receiver's posterior is highly dependent on which signal they observe, the expectation accounts for the signaler's probability of sending each signal given the true world:

$$\mathbb{E}[KL(P||Q)] = \mathbb{E}[\sum_{mind_{cg} \in M} P(mind_{cg}) \log \frac{P(mind_{cg})}{Q(mind_{cg})}] \tag{3.9}$$

A larger KL divergence occurs when the receiver is more uncertain about the contents of the boxes.

3.5.2 Results

When there is no wall, both models achieve the upper bound of possible performance. Consistent with our hypothesis, these models make different predictions when there is higher uncertainty in the world from adding a wall. When the wall is added, performance drops for both models; however, multiple comparison tests show that PaCo outperforms RSA at each level of ax (all $p_{adj} < .05$ under Tukey's HSD) except when there are four axes ($p_{adj} = .074$) (see Figure 3.4). When the receiver has four axes, there are no constraints on the action space and thus, considering actions is not able to restrict signaling behavior. Because PaCo cooperators take into account the receiver's action space, a less capable agent requires less information to do its best, making PaCo predict that it is sometimes better to tell their partner exactly how to act.

Even more striking, PaCo uses fewer tokens than RSA to achieve a higher task utility when the uncertainty in the world is high (Figure 3.5). With this uncertainty, constraints of the action space help reduce which signals are reasonable to consider. However, when the shadow is shown, PaCo uses more tokens than RSA, seemingly *over* informing. Just as considering the receiver's capabilities in the Wall condition had a benefit, here it has a drawback. When PaCo recognizes their partner as more capable, it sends a longer, more cautious signal to ensure the signal is not misinterpreted. Here signals that maximally reduce uncertainty are preferred, even when a shorter signal can be understood with high

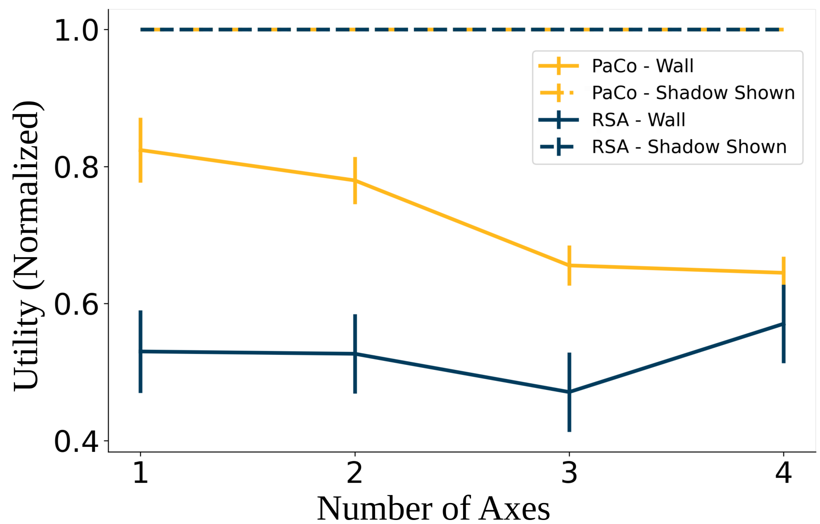


Figure 3.4: Utility achieved as a function of axes for RSA and PaCo with 95% CI. Dashed lines represent when there is no wall and solid lines represent cases when there is a wall. When there is no wall, both models achieve optimal utility 100% of the time. However PaCo is able to communicate more successfully when there is a wall and the uncertainty in the world is high.

probability. As the cost of communication is only .1, PaCo is willing to pay to be safe and ensure the receiver achieves the maximum reward possible. For example, if four of five boxes contain rewards, marking four boxes to open is semantically impossible to misinterpret, whereas marking one for avoid is possible (however unlikely) to be interpreted as a signal to open.

By definition, RSA always aims to provide the most informative message, whereas PaCo's action-based utility drives it to provide messages that maximize task outcome. From Figure 3.6, we see this clearly in the breakdown of model KL divergences. When the shadow is shown, both models always have virtually zero divergence, indicating that the signal can fully resolve the state of the world. However, in the Wall condition, higher uncertainty leads to a different pattern of results. RSA achieves a much smaller KL divergence than PaCo, indicating that RSA agents are likely to have a better understanding of the true world state,

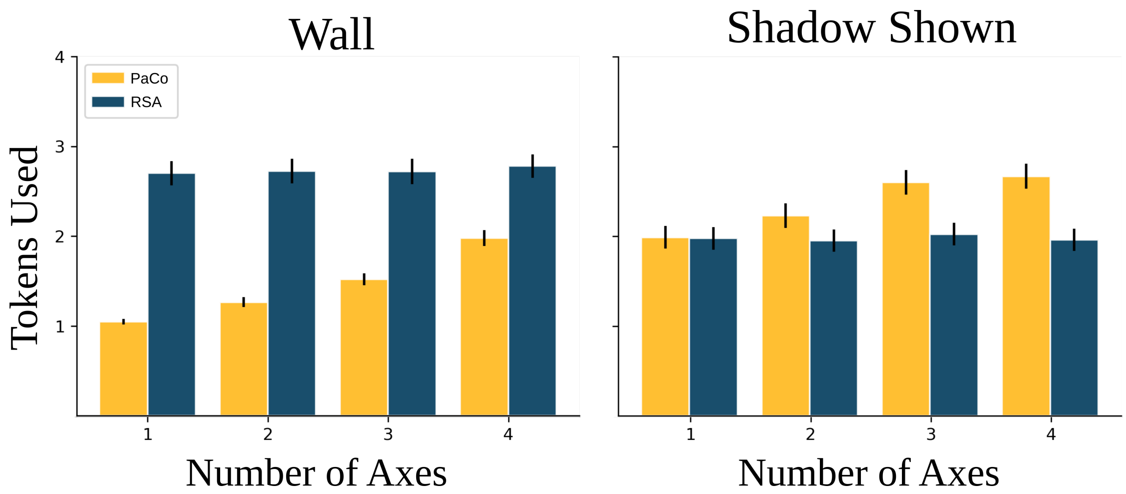


Figure 3.5: Proportion of tokens used by PaCo and RSA given the available receiver actions for cases with a wall (left) and when the shadow is present (right).

but that this alone is not enough to succeed at the task.

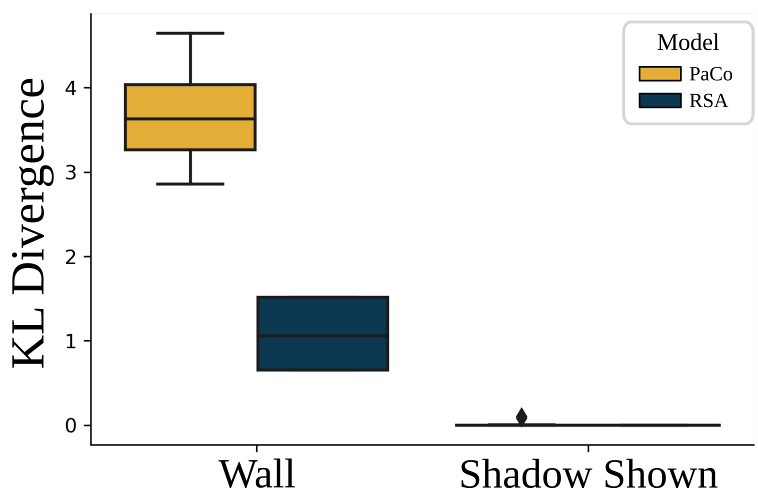


Figure 3.6: Distribution of KL divergence between receiver belief posterior and true world belief for N=2000 conditions from Simulation 2. True world belief distribution is adjusted to give non-target beliefs a small non-zero (1×10^{-6}) weight.

3.6 Simulation 3: Performance under Utility Driven Optimal Conditions

Coordinating minds and maximizing a utility calculus are two modeling components that do much of the heavy lifting in PaCo. Given this, we investigate how helpful deep recursion is for signaling in this task. We test whether the advantage of PaCo is present even without a complex partner model and to what extent adding recursion improves performance for PaCo and RSA. For both PaCo and RSA, we include players who do not model their partner. We then compare this to when each model has added levels of recursion to the receiver to see how the receiver having a model of their partner can change performance. In addition, previous simulations focus on specific environment cases that highlight features of how the models can emulate human behavior and interact with common ground. We also use this task to examine whether PaCo's advantage generalizes beyond the specific conditions from the previous simulations to cases on a larger scale and with few constraints in the environment.

3.6.1 Methods

3.6.1.1 Task Specification

For both PaCo and RSA, we start with literal communicators (naive or level-0) who do not model their partner. We compare this to models with additional levels of recursion to the receiver and signaler to see how having a partner model can change performance. To scale up reasoning in the environment, we look at environments with three to six boxes and remove all free parameters from the models, focusing on how well the models can perform in general settings without any prior biases. To measure the best possible performance under uncertainty, partners greedily select the action or signal with the maximum expected utility. We put a uniform prior over a token's meaning and remove all signaling costs. We then uniformly sample from the space of possible worlds all possible worlds with three to six boxes which have at least one scorpion and one banana. The number of axes and tokens are sampled independently such that there is at least one and at most $n-1$ for each, given a world with n boxes. The presence of a wall is also sampled as a binary variable. Each model at each reasoning level has a total of 400 simulated trials.

3.6.2 Results

For both models, having a model of one's partner improves communication substantially (PaCo: $\overline{x}_{S1R1-S0R0} = .34, t(439) = -15.3, p < .001$, RSA: $\overline{x}_{S1R1-S0R0} = .14, t(439) = -5.3, p < .001$). Moreover, there is a distinctive improvement in performance when the PaCo signaler is not a 0-level reasoner. A level-0 signaler randomly samples both what a token means and a consistent signal given that meaning leading to poor performance. However, an intelligent PaCo signaler can improve on this by taking into account what actions a receiver can take. When the signaler is not at level-0, all PaCo communicator pairs perform better than RSA communicator pairs of equivalent recursion depth ($p_{adj} < .001$) but this is not true when the signaler is a level-0 reasoner (average difference between PaCo - RSA at different receiver levels of reasoning: $\overline{x}_{R0} = -0.066, p_{adj} < .001$; $\overline{x}_{R1} = -0.067, p_{adj} < .001$ $\overline{x}_{R2} = -0.011, p_{adj} = .64$), adjusting for multiple comparisons using the Benjamini-Hochberg criteria.

These results indicate that PaCo's success does not necessarily rely heavily on deep recursion. Instead sensitivity to other task-related information may shift some of the burden off complex reasoning. Here, PaCo's flexibility in conveying information about actions and not just beliefs about the environment allow it to outperform RSA, especially in the absence of common ground information.

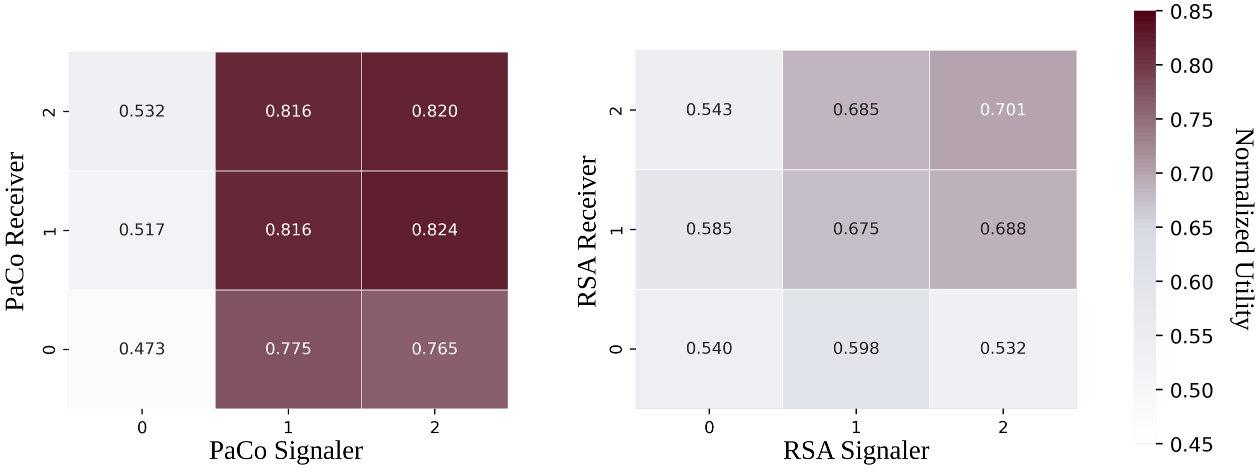


Figure 3.7: Utility achieved as a percent from optimal by communicator pairs at different depths of reasoning for PaCo (Left) and RSA (Right). Darker red indicates better task performance and higher signaler/receiver level indicates deeper recursive reasoning.

3.7 General Discussion

PaCo proposes a structured way to process common ground information to help resolve ambiguity in signaling. To disambiguate signals, we integrated ToM into RSA's pragmatic reasoning framework. This provides a holistic view of the interplay between common ground, the mind, and the shared environment which allows communicators to reason beyond referential signals about beliefs. PaCo also uses predicted action outcomes to determine the value of a signal, allowing us to argue for communication as a way to align cooperators' minds. Through modeling Misyak et al. (2016) as a case study, we highlighted (1) the importance of treating common ground as a multi-faceted constraint to signaling, which requires treating partners as rational and capable of achieving things in the instrumental world and (2) the benefit of framing communication as a means to coordinate perspectives, which highlights how different components of cooperators' minds interact to reduce reliance on deep social recursion.

3.7.1 Theory of Mind is Important for Processing Common Ground

Context is not one-dimensional. Different sources of uncertainty contribute information to the common ground: in this task, restrictions on world beliefs and available actions led to different human signaling predictions. RSA chooses signals based on how informative they are: signal utility comes from evaluating reduction of the receiver's uncertainty about the world. In contrast, PaCo considers rational action consequences and their underlying beliefs in conjunction, through its ToM mechanism. This ToM mechanism has demonstrated benefits in this task.

Achieved task utility in Simulation 2 established the theoretical contribution of PaCo's action-driven model over RSA's belief-driven one. Furthermore, taking action restrictions into consideration is important even outside of the highlighted conditions from Simulation 2. PaCo's performance advantage generalizes well in this task as demonstrated by Simulation 3. PaCo reached a higher asymptotic performance under optimal, maximal rationality. Furthermore, it did this at a shallow level of recursive partner reasoning and across different sized environments without relying on informative signal meaning priors or costs.

PaCo uses common ground information to both excel at the task and signal flexibly. In Simulation 1, the PaCo signaler used the same signal in different ways depending on the common ground information of number of signals, number of axes, and whether there was a wall. These results demonstrated a flexible fit to human behavioral data robust to a wide range of hyper-parameters.

In addition, we also saw further interesting results that can motivate future research. Two predominant strategies for communication emerged from the modeling work in Simulation 2's highlighted conditions. The first strategy was belief-driven (where bananas/scorpions are) and the second was action-driven (which boxes to open). While RSA is designed as exclusively belief-driven, our model naturally switched between these strategies based only on the calculation of expected outcome utility. Evidence for this came from looking at cases where the common ground information was highly uncertain in two specific analyses of Simulation 2. First, token usage as a function of available actions demonstrated that PaCo was able to say less while performing better at the task. Second, a PaCo receiver's beliefs diverged from the true world much more than RSA's. In those cases, the receiver did not have a good understanding of the world state, but still performed well because it interpreted signals as action requests.

This finding has real world implications, reflecting an interesting dynamic seen in everyday exchange. For example, consider the scenario where your kitchen sink gets clogged while your housemate is out running errands. Here, you must decide whether to inform your housemate of the situation — "the kitchen sink is clogged" — or tell them what to do — "pick up a plumbing snake on your way home." Which choice you make depends on how capable your housemate is; if they're very capable you may just inform them and let them figure it out how to fix it, but if they're not, it can be more useful to directly tell them what to do. This example highlights the trade-off in this "need to know" phenomena between action capabilities and what kind of information to provide your partner to achieve the desired result.

Signals are designed to change beliefs; however, using a ToM infrastructure puts understanding signals into a broader context. The meaning of a signal depends not only an agent's beliefs, but also the expectation that agents should do rational things based on those beliefs. This mechanism has contributed to PaCo's task utility and allowed PaCo to exhibit flexible signaling strategy switching sensitive to a rational agent's capabilities. The underlying support of a full ToM model provides solid grounding for communicators to coordinate with each other.

3.7.2 Communication is a Coordination of Minds

Grounding PaCo in a fully developed ToM still begs the question of how communicators are able to understand each other. Typically, recursive reasoning is emphasized as the way to model partner beliefs. However, when a joint inference over different mind components becomes the focus of signal resolution, a richer social phenomenon emerges. By jointly considering interacting mental states, communication can be viewed as a meeting of rational minds to help align perspectives. We examined the contribution of recursive reasoning to solve this task in Simulation 3, where PaCo was able to achieve a high performance with only a simple partner reasoning model. This indicates that some of the inferential burden associated with deep recursive reasoning can be lifted by adopting the perspective that communication is about coordinating our minds.

Without considering actions, there is no room for coordination. Under PaCo, a speaker considers actions by predicting their partner's reaction using the common ground. The speaker then evaluates those predictions using their private mind which may contain additional information related to the task. For a receiver, shared public knowledge provides necessary context to constrain the scope of possible ambiguous signal meanings. Under this approach, communication is treated as a means to update common ground information that the mind is build upon. Joint uncertainty over different components of the mind allows PaCo to naturally make trade-offs between competing components and decide which component is is most important to talk about in the moment as seen in Simulation 2 where PaCo sent informative signals as well as imperative signals depending on context.

3.7.3 Future Directions

We highlight the importance of actions, but planning in this task is limited in two ways: (1) planning over actions is a simple action cost calculus and (2) a cooperative shared agency perspective is ignored. In the original study, communicative roles are vastly simplified and the action space is naive. Here, fixed roles were assigned to the signaler, who cannot act, and receiver, who can. While having a predefined signaler and listener is common in signaling paradigms (e.g. Frank & Goodman, 2012; Selten & Warglien, 2007; De Ruiter, Noordzij, Newman-Norlund, Hagoort, & Toni, 2007), task oriented communication in the natural world is usually done between partners with overlapping abilities. Furthermore, planning was not highlighted, unlike many interesting sequential game settings studied in AI (Tambe, 1997; Grosz & Kraus, 1996). Moving away from one-shot actions opens up important cooperative considerations across time and space and brings signaling closer to conversation rather than a one-way exchange.

Directly telling a less capable agent what to do helped PaCo cooperators achieve a higher task utility efficiently; however, this modeling advantage diminished as the agent you talk to became more capable, with fewer action restrictions to narrow a signal's meaning. In real life, our cooperators are not often omnipotent, especially when interacting in the physical world. Instead, more nuanced utility considerations can shrink the space of affordable actions which puts strong constraints on how a signal should be interpreted.

While this task was constrained to arbitrary capability manipulations, agents could instead be constrained by cooperative rules such as joint utility dynamics in a spatial environment with barriers. Actions are still possible, but become expensive when considering joint utility. If the speaker always held the expectation that the receiver should do everything, this could actually increase the uncertainty of how to interpret an overloaded signal. Instead, working under a cooperative framework can actually enhance the efficiency of communication which has been tested in other work (Stacy et al., 2021).

In addition, framing planning as a joint problem between cooperative agents can actually help further constrain affordable actions and, as a result, reasonable signals. In these cases, it can be helpful to focus on a shared agency model of cooperative communication. PaCo can be extended to a shared agency model if we treat the minds that are coordinating as "We" minds, full of joint beliefs, joint desires, and joint intentions which produce rational joint action outcomes. Taking a joint stance has seen support in philosophy and cognitive science (Nagel, 1989; Tomasello, 2010; Gallotti & Frith, 2013) and in non-communicative cooperative modeling tasks (Tang et al., 2020; Kleiman-Weiner et al., 2016; Wu et al., 2021).

In this study, we've proposed a model to fit human performance from Misyak et al.

(2016) alongside many theoretical predictions. In more challenging cases simulated by the model, it is not clear whether people could actually act using these strategies. The results we find here, namely supporting shifting between communication strategies and replacing some of recursion with more nuanced partner predictions can hopefully serve as a starting point for future behavioral research. By answering these questions, we can build a better understanding of the mechanisms underlying coordination in communication.

CHAPTER 4

The Imagined We: Modeling Cooperative Communication to Coordinate Perspectives

One basic principle of traditional machine communication is to treat communication as an encoding and decoding process where each signal is uniquely mapped using a codebook. In reality, human communication is highly overloaded: a single signal can map onto many meanings. In this work we develop a formal model of this reasoning process, inspired by insights from both artificial intelligence (AI) and cognitive science. Our model views communication as a utility optimizing process for both understanding and influencing the minds and actions of other agents. A distinctive feature of our approach, which we call the Imagined We (IW) model, is the focus on shared agency instead of individual reasoning to capture the cooperative nature of communication. Under IW, each individual agent models a We mind, simulating a centrally controlled super-agent. While the We mind is not real, agents act as if it exists; thus communication makes information public and shared to help coordinate different simulated versions of We. The current study models IW in a set of simulated signaling tasks that incorporate linguistic ambiguity in gridworlds. We show IW is capable of successfully signaling meaning under high ambiguity. Additionally, when signaling becomes costly, IW maintains task performance while substantively decreasing the amount of information exchange required for successful communication. When agents recursively model each other, IW outperforms baseline models of pragmatic reasoning without shared agency at even the shallowest level of recursive reasoning. These results highlight how constraints from rationality, shared knowledge, and cooperative logic can do much of the heavy-lifting in communication.

4.1 Background

Imagine Audrey and George are sitting at a table with a glass of water precariously close to the edge. Audrey exclaims "the glass!" and George instantly moves his glass away from the table's edge. Here, "glass" is sparse, leaving the listener infer why the glass is relevant and how to best respond. Additionally, "the glass" is overloaded in what is refers to. It might refer to a broken window pane in one context, or the request for a refill in the next. These simple, everyday exchanges involve spontaneity and indirectness, capabilities that are the hallmark of intelligent "inference-making machines" (Sacks, 1985). When choosing what to say, humans rely on what they and (their estimate of) their partner know, see, want, and want to do. The way humans communicate is different from how models of communication, such as those that build on information theoretic principles, are typically constructed. Humans are highly sensitive to context, both linguistic and non-linguistic, which allow them to resolve the ambiguity in sparse exchanges. The importance of including context in models of communication has long been recognized but has been notably difficult to model (Sperber & Wilson, 1986; Levinson, 1983). Recently, works have begun to formalize aspects outside of the current exchange of signals themselves, such as a history of past experiences (R. X. Hawkins et al., 2017), common ground expectations (Bohn, Tessler, & Frank, 2019), and the space of possible referents itself (Ashok Kumar, Garg, & Hawkins, 2021).

Leveraging computational work and empirical advances from artificial intelligence (AI) and cognitive science, we develop a human-like model of cooperative communication. This model contains three novel properties:

1. It views communication as an action to change others' minds, which can be formalized by a model of agency ("glass" is designed to make you believe the glass is important

and act to save it).

2. This model of agency goes beyond the individual, instead taking on a shared We perspective capable of creating joint plans given joint beliefs to maximize joint utility. Here "jointness" acts as a constraint on communicative intent: signals are constrained to interpretations that are expected to improve the joint utility under a cooperative logic where agents must treat partners with respect and commitment to find jointly efficient solutions (signaling "glass" implies the message is relevant to the listener George; thus "glass" is not the one closer to Audrey – that would be Audrey's responsibility to keep safe).

3. When considering joint planning and joint utility, non-linguistic context such as physical and visual cues are included ("the glass" is not about the linguistic properties of the signal; it refers to the glass's physical instability).

4.2 Linguistic Context from Signaling Pragmatics

We supplement AI models of planning with ideas from cognitive science, to focus on two types of context in communication: context that comes from what you can say as well as context that comes from what you can do. First, we look at linguistic context. Referential language games are a typical setting in linguistics used to study overloaded communication (Wittgenstein, 1953). In these games, the environment contains a set of potential referents with features (e.g. shape, color). A listener aims to understand which referent a speaker is indicating from a potentially ambiguous signal. This has been more recently modeled in cognitive science using the Rational Speech Act (RSA) framework (Frank & Goodman, 2012; Goodman & Frank, 2016), a Bayesian model of signal pragmatics that uses the mutual assumption of communicative cooperation to resolve potential ambiguity. Here cooperative signaling involves being truthful, relevant, and straightforward with respect to the communicative goal of conveying a referent (Grice, 1975). This assumption constrains both signals and their interpretation as communicators must speak (and assume others are speaking) in line with this principle. In addition, assuming that the speaker has produced the best utterance possible allows a listener to discard interpretations that are literally correct but could have been referred to more directly. We introduce these ideas formally in Section 4.6.1.

RSA addresses ambiguity in communication by emphasizing the importance of linguistic context and has been successful at modeling phenomena such as metaphor (Kao et al., 2014), redundancy (Degen et al., 2020), and politeness (Yoon, Tessler, Goodman, & Frank, 2017). However, this model remains primarily focused on linguistic context, where pragmatics is largely determined by the available vocabulary – both what the speaker says and chooses not to say. These are key features for language but may not be enough to capture communication. In many real life situations, communicative context is more broadly defined to include tasks executed in the visual and physical environment.

4.3 Non-linguistic Context from Agency in the Environment

4.3.1 Theory of Mind as Agents Modeling Other Agents

The second integration of AI and cognitive science comes from the insight that a large part of non-referent context can be captured by considering not just what an intelligent agent should say but also what they want to do, why they want to do it, and how they can accomplish it. Here, we rely on an existing line of computational work which emphasizes agency and is capable of reasoning over rational actions in the physical world. This agency approach provides a means to connect ambiguity in signaling with non-linguistic context from planning in the shared environment to make up for the sparsity and ambiguity of communication.

An early agency-based approach heavily studied in AI to modeling human reasoning centers around describing the underlying mental states of an agent that drive its actions: namely beliefs, desires, and intentions (BDI). The BDI framework originated from a perspective in philosophy that regards beliefs as the informational states of the mind, desires as the motivational states of the mind, and intentions as the deliberative states of the mind (Bratman, 1987). This approach has garnered attention for decades (Adam & Gaudou, 2016; Georgeff, Pell, Pollack, Tambe, & Wooldridge, 1998; Norling, 2004; Rao & Georgeff, 1995); however, on its own, this framework cannot support humanlike communication. This is because BDI agents do not easily represent uncertainty in a probabilistic manner which, in turn, cannot support ambiguity in communication. While the BDI architecture has been extended to assess metrics such as truthfulness and responsiveness as cooperative heuristics in communication (Singh, Padgham, & Logan, 2016); again, these are rule-based decisions based on accumulated signals from other agents. Moreover, communication is not simply planning based on one's own beliefs, desires, and intentions as under BDI, but also involves inferring the contents of others' minds in order to send relevant signals.

Using the underlying representation of mental states that the BDI framework offers, Bayesian inference can be used to interpret signals. This interpretation is not based on a fixed mapping, instead it is a probabilistic inference of other agents' minds – including beliefs, desires, and intentions – and the rational planning process over them. This approach, stemming from Theory of mind (ToM) in cognitive science, is a widely established type of social reasoning (Gopnik & Meltzoff, 1997; Wellman, 1992) which predicts that individuals act rationally according to underlying mental states, as defined by the BDI model. This process has been formalized as a Bayesian inference problem in Bayesian Theory of Mind (BToM) (Baker et al., 2009; Baker & Tenenbaum, 2014). In forward planning, agents aim to maximize their utility according to their mental states while minimizing costs of acting in the world similar to the BDI formulation. By observing others' actions a, Bayes rule inverts the planning process to help infer the underlying mind containing beliefs, desires, or intentions that may have produced those actions given the environment w.

$$P(mind|a, w) \propto P(a|mind, w)P(mind|w) \tag{4.1}$$

While we focus on BToM as the foundation for modeling communication in interacting agents, it is only one example of a broader class of models which aim to capture how agents model other agents (Fan & Yen, 2004; Doshi, Gmytrasiewicz, & Durfee, 2020).

Interpreting a signal in isolation is often not enough to communicate successfully; instead, BToM allows communicators to model their partners' minds which are full of situational context critical to resolving ambiguity. The context offered by the mind has two parts. First, the contents of mental states define the space of beliefs, desires, and intentions and second, the rational planning process provides principles that guide how mental states interact to generate actions.

4.3.2 Communication as a Tool to Coordinate Plans

We treat communication as a planning problem driven by utility maximization based on BToM: able to reason over mental states and actions of others. From this perspective, communication is broader than language games from cognitive science which focus on guessing and changing the contents of others' minds without linking the mind to actions. Instead, we view communication as a problem that is highly consistent with a fundamental principle of AI: rational agents choose actions that maximize their utility (Russell, 2019). This principle is closely connected to multi-agent problems, such as interactive partially observable Markov Decision Processes (I-POMDPs) (Doshi et al., 2020; P. J. Gmytrasiewicz & Doshi, 2005), which involve agents reasoning about each other. Here, when agents have partial information about the world, communication can improve task performance. However, AI formulations do not highlight the flexible, instantaneous interpretation of a signal's meaning. Our goal is to take this perspective of understanding communication in the context of planning while maintaining the cognitive science focus on flexible meaning interpretation.

In an I-POMDP setting, signals can provide information that can change others' beliefs allowing agents to take good actions with respect to their personal reward functions. This captures the idea that communication should be understood in the context of the task, instead of focusing purely on linguistics. The interactive aspect of this formulation is that state changes, observations, and rewards are jointly determined by the decisions of all agents. A cooperative alternative, which has been extended to include communication, is the decentralized POMDP (Dec-POMDP) (Bernstein, Givan, Immerman, & Zilberstein, 2002). However, similar to their single agent counterparts (Crandall, 2020; Hayashi, Ruiken, Hasegawa, & Goerick, 2020), typical solutions to Dec-POMDPs execute plans in a decentralized manner but rely on centralized planning to learn and agree upon agents' policies a-priori (Oliehoek & Amato, 2016; Spaan, Gordon, & Vlassis, 2006; Xuan, Lesser, & Zilberstein, 2001) or by assuming a fixed policy of the signaler (P. Gmytrasiewicz, 2020). Unlike the language game, the goal of communication is to improve task performance by connecting signaling to actions in the physical world, the signal is an observation reasoned over according to predetermined knowledge about agents' policies. While approaches from these studies demonstrate the importance of understanding communication in the context of tasks and share the insight that signals can be tied to action utilities, they do not focus on the pragmatic aspects of communication inherent in RSA for flexible signal disambiguation.

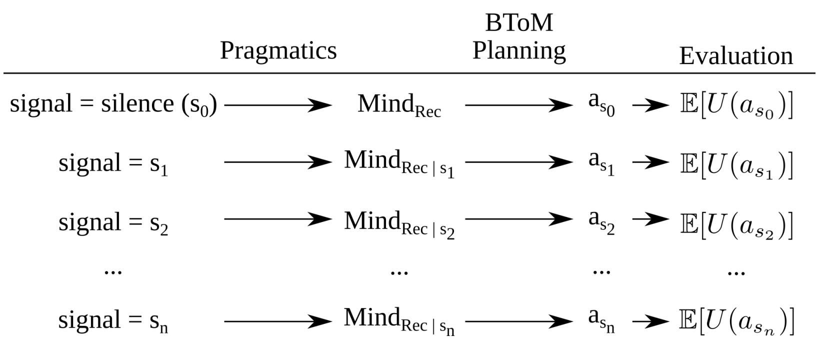


Figure 4.1: Connection between signals and actions through the mind.

The key is to combine RSA pragmatics with an agency model that plans over actions in a cooperative task. BToM provides a rich framework for harnessing intuitive action costs and preference rewards (Jara-Ettinger et al., 2016; Ullman, Spelke, Battaglia, & Tenenbaum, 2017), which allow for planners using the environment with intuitive physical constraints (e.g. agents cannot move through barriers). Moreover, framing signals as another type of rational action allows BToM to be readily integrated with communication to connect signals back to actions in the environment. In line with this, there have been two examples of recent efforts which treat communication as an action to improve planning in the task using individual BToM (Sumers et al., 2021; Jiang et al., 2021). First, work from Sumers et al. (2021) considers the trade-off between being maximally informative (belief-oriented) and maximizing task rewards (action-oriented) and their conjunction. This allows for fine-grained comparison of behavior under different models of signaling and demonstrates how changing others mind can help them understand a reward signal. Alongside this, a second example (Jiang et al., 2021) provides a model of relevance that relies on pragmatics and ToM to resolve an overloaded pointing gesture. Using a classic AI paradigm, Wumpus World (Yob, 1975), a pragmatic helper who knows the location of the Wumpus can choose to emphasize observations that a hunter receives by pointing. The hunter reasons that the helper aims to be relevant and interprets the pointing gesture in the manner that can best improve the hunter's utility, which requires an integration of BToM and pragmatic linguistic reasoning.

4.4 A Shared Agency Perspective on Cooperative Communication

Previous work has focused on individual ToM reasoning where agents reason individually about other agents; here we focus on cooperative communication which poses constraints on partners due to its shared nature. Specifically, we highlight the importance of cooperative logic when communicating while performing a joint task. That is, being cooperative is about being truthful, efficient, and easy to understand (Grice, 1975). We go one step farther arguing that it also involves treating others fairly as equals, meaning cooperators should search for jointly efficient solutions. Requests must be justified from a joint perspective: agents should not ask others to do things easier for themselves to achieve. For this reason, we argue it is important to model communication as not just an agency problem but rather a *shared* agency problem.

One school of thought in philosophy posits that under shared agency, collaborators must be committed to achieving a joint intention as a collective body (Gilbert, 2013). Under this perspective, agents individually reason over how someone viewing the task from a third-party perspective would coordinate actors (Tomasello, 2010) or "view from nowhere" (Nagel, 1989) in order to approximate shared agency. This could be viewed as a centralized "We" agent who coordinates agents as if they were limbs of a body. The problem with this centralized We is that it does not exist in reality: only You and I exist as individuals. That is, each agent must separately imagine We. While the aim is to model the same We, in reality each agent may imagine a slightly different version. When these versions of IW are synchronized and represent the same contents, agents can coordinate smoothly.

Even without explicit communication, shared agency can put strong enough cooperative constraints on actions to allow for successful coordination. The plural We stance has been modeled in AI settings (Grosz & Hunsberger, 2006; Jennings, 1995; Tambe, 1997; Grosz & Kraus, 1996; Levesque et al., 1990), typically guided by heuristics and rule-based reasoning. In addition, more recent computational successes adopt a Bayesian framework and extend BToM to a shared problem. Adopting a shared BToM perspective is naturally sensitive to action context and makes planning explicitly joint. Planning over the joint space of all cooperators' actions who share a utility function has been shown to produce successful cooperation in a physical coordination task (Kleiman-Weiner et al., 2016) and cooking paradigm (Wu et al., 2021). Moreover, a stronger formulation of We, achieved through explicit representation of joint intentions — in addition to shared rewards and joint actions — has allowed agents to robustly bootstrap joint commitment to one of many equivalent, arbitrary goals in a hunting paradigm (Tang et al., 2020).

This formulation mirrors individual BToM where the mind is narrowly defined in terms of its underlying components: beliefs, desires, and intentions. In this case the mind is shared and contains joint beliefs (b_{we}), joint desires (d_{we}), and joint intentions or goals (g_{we}).

$$P(mind_{we}) = p(b_{we})p(d_{we})p(g_{we}|b_{we}, d_{we}) \tag{4.2}$$

It is notable that none of these modeling works involve communication; instead, agents can spontaneously coordinate using constraints from joint planing while reasoning from a shared agency perspective. This has further implications for how communication can be modeled. Adding even a sparse signal — a nod, shared glance, or pointing gesture — on top of a shared agency foundation can further improve intelligent interpretation under increasingly flexible situations. This gap between models of communication and cooperation can be bridged with the insight that communication can be viewed as a social tool to enable increasingly intelligent cooperation (Tomasello, 2010). Thus models of communication should be built in the context of cooperation which involves signaling, reasoning, and acting with others – unlike purely linguistic formulations or fixed signal mappings. To reflect that each individual represents the collective shared group of cooperators, we call our communication modeling framework the Imagined We (IW).

Modeling shared agency in the context of cooperation can provide a solution to a long-standing challenge in multi-agent interaction: recursion. Traditionally, models depend on recursive partner reasoning where cooperators must make a decision about which layer of recursion to stop at (Camerer et al., 2004). IW makes inference a one-way process. This relieves some dependence on this recursion: an agent reasons about IW but IW does not need to reason about individuals. In addition, communication is public and transparent. Shared knowledge serves as the basis for interpretation of signals whereas private knowledge is excluded from consideration (Clark & Brennan, 1991; Clark & Marshall, 1981). Thus, an interpretation of an overloaded signal occurs in the context of mutual transparency: a signaler expects a receiver to resolve overloading using the existing mutually shared knowledge. Moreover, the signals discussed in this work are overt, carrying an additional layer of communicative intention (not only do I want you to know X, I also want you to know that I want you to know X), enforcing norms of helpfulness or cooperation (Tomasello, 2010). The role of communication, then, is to make information public to align different versions of IW. Signals in conjunction with ToM reasoning offer an especially versatile and flexible mechanism to do this that allow even young toddlers to employ cooperative logic to constrain the meaning of an ambiguous requests for help (Grosse et al., 2010).

4.5 Task: Cooperative Referential Signaling in Gridworld

We start by introducing our modeling task then provide computational formalism, using the task as a running example. We test IW in an gridworld task to demonstrate its ability to communicate successfully in a cooperative setting. This task combines feature overloading, which demands the language pragmatics studied by frank2012predicting, but is enriched by a spatial scene that requires joint planning similar to the overloaded helping from grosse2010.

In this task, a signaler and a receiver cooperate to reach a target item among a set of items placed in a gridworld environment. Each item has multiple feature dimensions that can take on different values. In Experiments 1 and 2, these are shape (circle, triangle, square) and color (orange, green, purple). We start with the case where signaling is costless but restricted to a single feature of the target (e.g. "purple"), adding ambiguity and increasing the chance it will refer to more than one item in the environment. In Experiment 3, we relax this assumption and allow signals to contain multiple features and look at items with up to 5 feature dimensions. However, signalers are still motivated to send short signals as sending each additional feature comes at an added cost.

In play, the signaler acts first – she may walk to and select an item (incurring the appropriate action cost), send a signal to her partner (free), or quit the trial (earning a utility of zero). If the first agent sends a signal, the receiver then gets a turn. The trial ends after the receiver's turn or the target item is reached, whichever comes first.

Cooperation occurs in a nearly-fully observable setting. The key is the only bit of asymmetry of information between agents: only the signaler knows which item is the target. If either agent reaches the target, both receive a reward (+8); however, each step taken by either agent incurs a cost (-1) which is shared by both agents. We calibrate the setup of the environment so that the signaler is motivated to ask for help: if the signaler acts for herself when it is better to ask for help, the expected utility is around zero. However communication comes with a risk as it is often overloaded. As a result, if the receiver misinterprets the signal, he may go to the wrong goal and incur a large action cost without gaining a reward. In cases where the cost of the signaler walking to the goal herself is too high, and the signaler cannot communicate clearly, the rational decision may be to quit.

In each trial, the set of shapes in the environment is randomly sampled and located with the constraint that the target is closer to the receiver. As a result, the optimal utility can only be achieved when the signaler successfully asks for help which gives us a way to see the difference between communication models. In addition, to make joint planning more interesting, we add a physical barrier which is either near the receiver or near the signaler (see Figure 4.2). The closer the barrier is located to the signaler, the more ambiguity the receiver needs to handle from a joint action perspective. This is because a larger portion of the environment becomes the receiver's responsibility which makes the constraints from joint utility less likely to be useful for understanding a signal. However, pragmatics which focus on linguistic features should not be impacted much by this barrier.

4.6 Computational Modeling

IW integrates two types of context: linguistic pragmatics and joint efficiency of actions. Thus, we compare IW to two baselines reflecting these individual components and two distinct lines of reasoning. In addition, we make a comparison to the central control optimal solution (CC). CC is a true, not imagined, central controller which reflects how the two agents would rationally coordinate with perfect information: the ceiling of achievable utility. This is calculated with value iteration over the concatenation of the individual agents' action spaces.

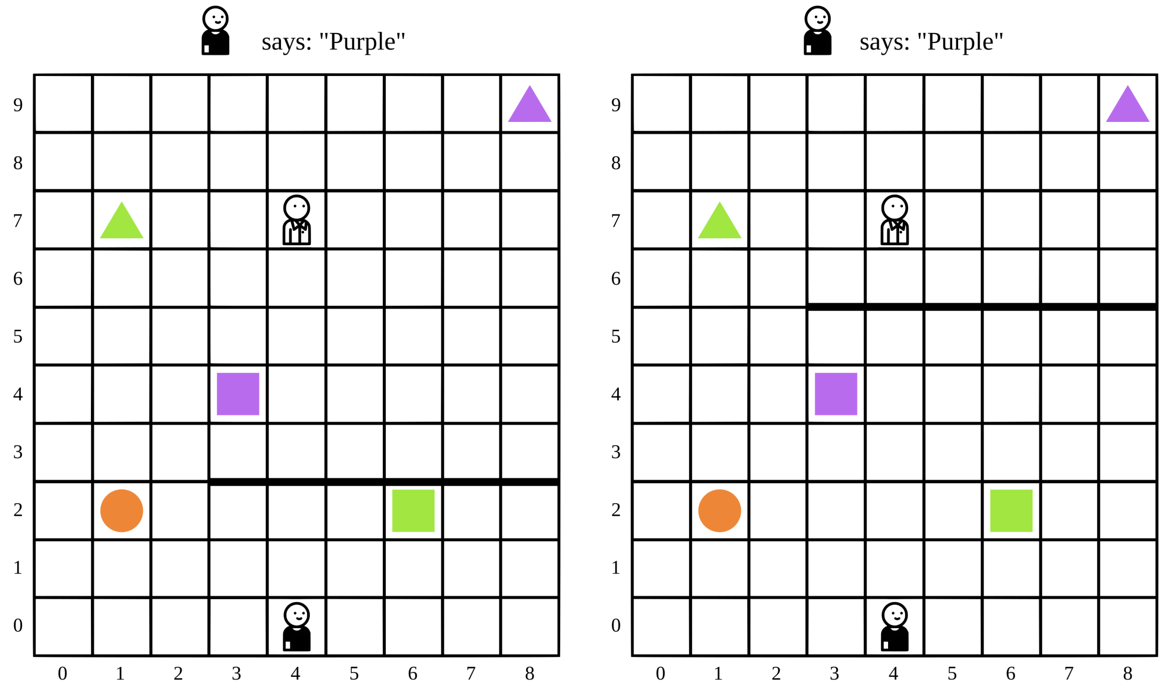


Figure 4.2: Example trial setup with a barrier near the signaler on the left and barrier near the receiver on the right. Moving the barrier up toward the receiver changes the joint utility dynamics. Both purple items are equivalent from the receiver's individual perspective, but the purple triangle is jointly efficient.

4.6.1 Modeling Linguistic Pragmatics: RSA

Here we provide the computational formalization of RSA which has been used to model linguistic ambiguity and explain an extension to accommodate this task. This formulation acts as one of the components of IW reasoning as well as a baseline comparison to IW in this work. The underlying aim of a RSA speaker (sp) is to describe a state of the world ($state \in \Omega_w$) with vocabulary ($signal \in \Omega_{sig}$) that is true but may have multiple possible referents. In the task we introduced here (see Figure 4.2 for example), the state represents the true target item (e.g. purple circle) and the vocabulary is limited to features of these items (e.g. "purple"). This decision making process is based on the signaler's estimation of the utility of a signal when describing *state* (saying purple to describe the purple circle). Signalers noisily maximize their expected utility (soft-max), leading them to send approximately rational signals, where $\beta \in [0, \infty)$ represents the degree of rationality.

$$P_{sp}(signal|state) \propto e^{\beta \mathbb{E}[U(signal, state)]} \tag{4.3}$$

In the typical language game setting, the speaker's utility is based on her estimate of the listener's probability of correctly identifying the target.

$$\mathbb{E}[U(signal, state)] = p_{Listener}(state|signal) \tag{4.4}$$

Calculating this utility requires a model of the listener. The simplest possibility is to assume the listener interprets the signal literally; as a result, any item with the signaled feature is equally likely to be interpreted as the target. However, the speaker may believe that the listener is pragmatic and also reasons about the speaker in order to make decisions. This becomes a recursive process where a more sophisticated listener model depends on a more sophisticated speaker model and vice versa. A pragmatic listener (lp) uses Bayesian inference to interpret a signal based on a model of a simple speaker.

$$p_{lp}(state|signal) \propto p_{sl}(signal|state)p(state) \tag{4.5}$$

As an entering point for recursion, this speaker model is literal (sl) and uniformly samples truthful signals — features that belong to the target item — without consideration for the listener. However, this recursion can be built up in additional layers for increasingly complex partner models.

In its original formulation, RSA is intended to be a language only model; however, the current task allows the speaker to make non-communicative choices. Here, the speaker has the additional choice of whether they would like to quit or perform an action, walking to an item instead of sending a signal (DIY). To make RSA a fair baseline, we build a mechanism for considering action utilities on top of the traditional communication component. We call this adaptation action RSA (aRSA). Instead of selecting only a signal, an aRSA speaker now makes a utility driven decision based on $d \in \{DIY, quit, \Omega_{sig}\}$ instead of *signal* according to Equation 4.3.

Now the expected utility of a signal is not only the probability of the listener understanding, but also considers the consequence of the listener's beliefs in terms of action utility. Using the traditional communication model to generate $p_{lp}(w|signal)$, the listener's interpretation, the signaler evaluates the receiver's utility: the action consequence of that interpretation. This occurs via value iteration as the listener's cost of traveling to the item a_r plus whether a_r yields a reward: $\mathbb{1}[a_r = goal_t]$.

$$\mathbb{E}[U(signal, goal_t)] = \sum_{w \in \Omega_w} p_{lp}(w|signal) U(a_{pl} = w, goal_t) \tag{4.6}$$

This is different from IW's formulation of utility as it still lacks "jointness" – that is, action utility helps a signaler decide whether to communicate or act, but once a signal is sent, the receiver uses standard RSA pragmatic reasoning to decide what that signal means.

4.6.2 Modeling Jointly Efficient Actions: Joint Utility

The second important component of IW and another baseline comparison focuses on joint utility. Joint Utility (JU) agents share a joint perspective on "who should do what." JU signalers always signal when it is jointly efficient for the receiver to obtain the target, regardless of how difficult it is for the listener to interpret. At the top level, a JU signaler samples a decision $d \in \{DIY, signal, quit\}$ according to Equation 4.3 and similar to aRSA. However, here the decision of whether to signal is made first and based on the optimal action the receiver could take in the environment.

$$\mathbb{E}[U(d, goal_t)] = \begin{cases} U(a_{sp} \rightarrow goal_t) & \text{if } d = DIY \\ U(a_{lp} \rightarrow goal_t) & \text{if } d = signal \\ 0 & \text{if } d = quit \end{cases} \tag{4.7}$$

In the cases where $d = signal$ is sampled at this first step, the actual selected signal is then sampled from all literally true signals consistent with $goal_t$. Similarly, a JU receiver, upon hearing a signal, constrains herself to consistent interpretations, weighing those interpretations according to their joint utilities.

4.6.3 Modeling Signaling under the Imagined We

A speaker cannot send a signal just because it is more convenient for the receiver to take an action without regard for how the receiver will understand the signal. At the same time, the speaker cannot send a signal solely based on linguistic properties without regard to its potential consequences. IW takes both the linguistic pragmatics of RSA and joint planning of JU and integrates them to resolve signal ambiguity. As a result, actions become both a way to evaluate the consequence of communication and a constraint for how to disambiguate signal ambiguity.

In the present study, we focus on the case where the environment is fully observable so there is no uncertainty in beliefs and the task is specified for all agents so that there is no uncertainty in desires. Instead, the uncertainty lies in which goal is the target $goal_t$, which only the speaker privately knows. The speaker is approximately rational and selects a signal according to its utilities as in RSA:

$$P(signal|goal_t) \propto e^{\beta\mathbb{E}[U(signal, goal_t)]} \tag{4.8}$$

Unlike traditional RSA, IW defines the utility of a signal by looking at the utility of the outcome actions under the task, weighted by how often those actions are expected to occur:

$$\mathbb{E}[U(signal, goal_t)] = \mathbb{E}_{P(a|signal)}[U(a, goal_t)] \tag{4.9}$$

This framework actually serves to coordinate different perspectives: (1) Action prediction: The speaker predicts how a signal can change the IW mind (here, shared goal: $goal_{we}$). (2) Action evaluation: The speaker evaluates how good that change is according to their private knowledge of $goal_t$. The evaluation of $U(a, goal_t)$ includes the cost of taking a and the reward if a achieves $goal_t$.

Action prediction can be further broken down by connecting signals to actions via the mind. First, signals change the IW mind, making some goals more likely than others. Second, using the BToM likelihood function for action planning, we can calculate which actions are rational conditional on a given joint mind. We assume actions are conditionally independent from signals given the mind, captured by the intuition that signals can only influence actions through the mind:

$$P(a|signal) = \sum_{goal_{we}} P(goal_{we}|signal)P(a|goal_{we}) \tag{4.10}$$

Traditional ToM planning yields $P(a|goal_{we})$ and Bayesian inference allows us to measure how observing a signal will change the distribution of inferred goals. For the likelihood function we use a measure of consistency (Is this message truthful given the goal?), similar the literal speaker from RSA:

$$P(goal_{we}|signal) \propto P(signal|goal_{we})P(goal_{we}) \tag{4.11}$$

IW is a shared agency account of modeling language that is able to integrate different types of relevance — language pragmatics and intuitive utilities — to communicate rationally under different sources of ambiguity. See Appendix A for full pseudo-code.

4.7 Simulation 1: Amount of Ambiguity

In this experiment we demonstrate how different models communicate as a function of the degree of overloading. We achieve this by manipulating the number of items in the environment from two to nine and comparing this across two barrier positions. We randomly sample items without replacement out of the set of nine possibilities (3 colors by 3 shapes). As the signaler can only communicate a single feature (a color or shape), increasing the number of items increases the potential overloading in the environment. When there are only two items, it is always possible to uniquely identify the target; however, when there are three to five items, the signal is potentially overloaded. At six or more items, the feature is guaranteed to be overloaded. As the goal of successful communication is to efficiently reach the target, we compare the utilities achieved by each model when faced with the exact same scenario. In addition, we look at the breakdown of signaler decisions (DIY, quit, signal) and receiver decisions (correct/incorrect interpretation of signal).

4.7.1 Results

Figure 4.3 shows the utilities achieved by each model, as ambiguity increases. Different environments have varying maximum achievable utilities which can be obtained by the CC. Thus, we measure task performance of each model as the percent from the maximum possible determined by CC.

First we focus on the case where the barrier is located near the receiver. When compared to the utility of always doing it for yourself (DIY), all models of communication lead to substantial gains in utility (mean difference (95% CI): $\overline{x}_{IW-DIY} = 3.78(3.68, 3.88)$, $\overline{x}_{RSA-DIY} = 2.02(1.90, 2.15)$, $\overline{x}_{JU-DIY} = 2.38(2.22, 2.55)$). Benjamini-Hochberg adjustments for multiple comparisons to paired t-tests find all $p_{adj} < .001$. However, these models' capacity to overcome ambiguity varies: the advantage disappears at high levels of ambiguity for aRSA and JU, but not for IW. Across any number of items in the environment, IW outperforms other baselines and only begins to deviate from the CC model when the uncertainty is very high. At the highest level of ambiguity (9 items) IW achieves 71.5% (CI: 64.6–78.4%) of the optimal utility on average, while aRSA achieves 3.7% (CI: 0.5–6.8%) and JU achieves 19.4% (CI: 4.4–34.4%). This demonstrates how communication understanding is significantly enhanced by the integration of linguistic and action-based reasoning in this task. When the barrier is located near the signaler, an extremely similar pattern emerges; however, this condition is harder for JU and IW. Both models perform systematically higher when the barrier is near the receiver than near the signaler. This is not the case for aRSA (see Figure 4.3).

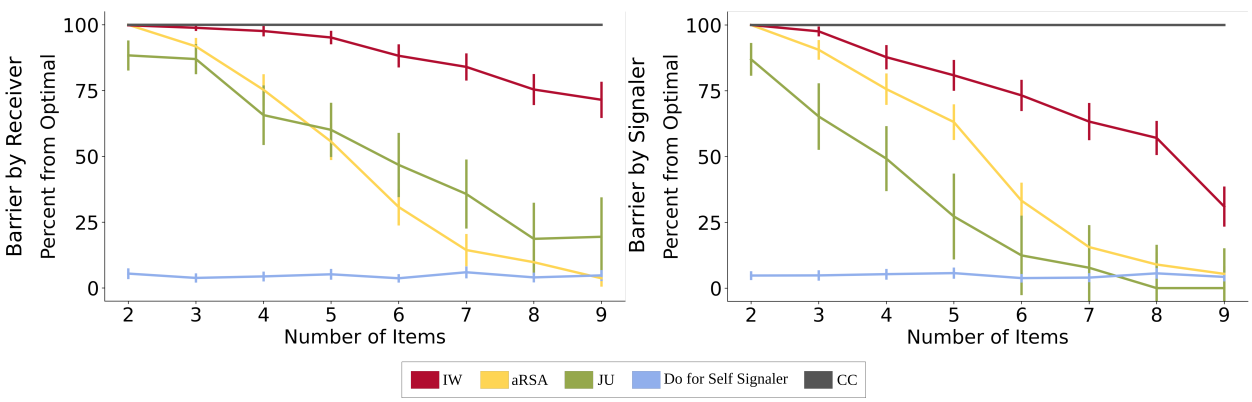


Figure 4.3: Achieved utility (95% CI) measured as the percent from optimal for each model under varying degrees of ambiguity. (Left) When barrier is near the receiver. (Right) When barrier is near the signaler. N=2000 trials per model; $\beta = 4$ for all models.

We make a more fine-grained comparison between models to understand what contributes to differences in achieved utilities by breaking down signaler and receiver decisions. As ambiguity increases, model decisions diverge hugely (Figure 4.4). As expected, the JU model always communicates without considering whether the receiver will understand the signal. As a result, the receiver often incorrectly interprets the signal, which dramatically increases errors as the number of items increases. In aRSA, the trend is dramatically different. aRSA signalers consider how their partner will interpret the signal; thus the percentage of misunderstanding is relatively low. However, only using features often leads to large uncertainty in the receiver's interpretation so the signaler is rarely willing to communicate, almost exclusively favoring quit by nine items. IW is able to combine these reasoning strategies to perform well across all levels of ambiguity: there is a much smaller decline in successful communication when the uncertainty is large. Similar to JU, IW has a strong preference for communication; furthermore, similar to aRSA, IW considers how that communication will be interpreted. As a result, even at nine items communication is highly successful and mostly avoids the pitfalls exhibited by JU and aRSA. Collectively, these results demonstrate that IW is capable of handling overloading to successfully communicate even under high signal ambiguity. Moreover, reasoning over both vocabulary and actions are vital to this disambiguation process.

4.8 Simulation 2: Level of Recursion

One strength of IW is that integrating the additional constraints from cooperative joint planning can often quickly resolve ambiguity, lessening reliance on deep recursion. This may provide a novel answer to why everyday pragmatic language often feels quick and easy. Here we demonstrate this by looking at how performance changes as a function of deeper reasoning for IW and aRSA at different levels. We focus only on the contrast between IW and aRSA as JU does not have a model of one's partner so recursion does not make sense. As introduced previously, both IW and aRSA do utilize recursion to interpret signals pragmatically: the signaler models a receiver to determine which signals will be understood and the receiver models the signaler to interpret a signal. The only difference is that in RSA the target of inference is your partner's beliefs whereas the target of inference in IW is the shared IW mind (the part where the signaler takes the receiver model to make an inference about We can be shown in line 5 of the signaler algorithm in Appendix A). We compare the utilities achieved by different reasoning levels of speaker and receiver playing this task. In IW, a joint

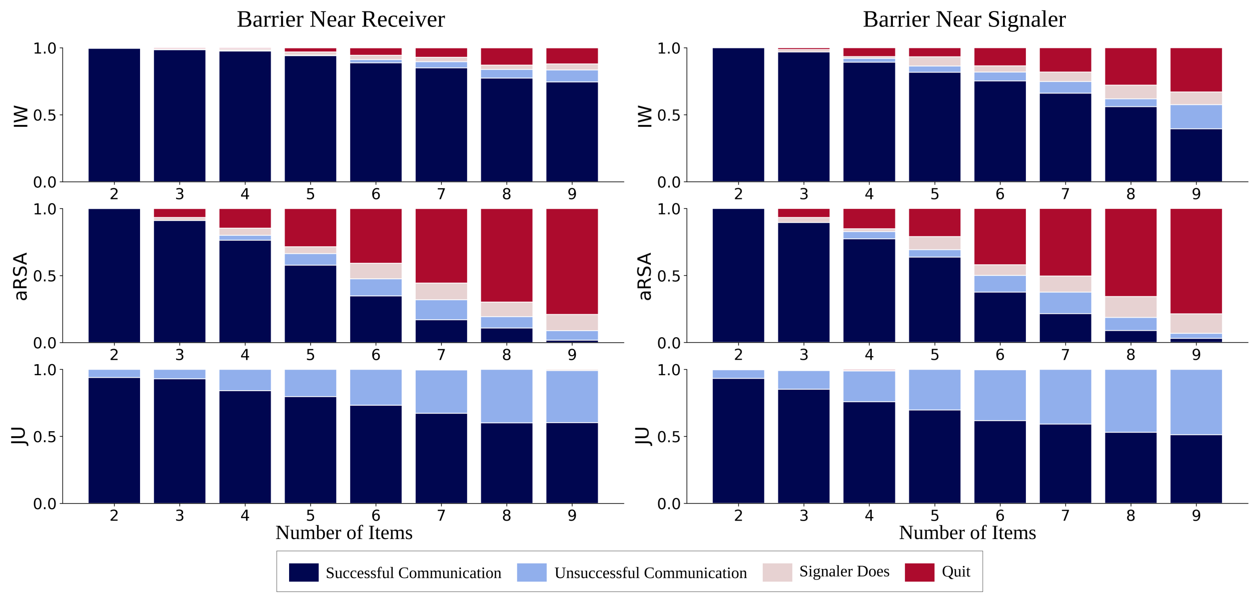


Figure 4.4: Breakdown of agent decision (as a proportion) for each model under varying levels of ambiguity. The decisions are (a) Successful communication: the signaler communicates and the receiver goes to the correct goal (b) Unsuccessful communication: the signaler communicates and the receiver fails to choose the correct goal (c) Signaler does: the signaler forgoes communication and walks to the target (d) Quit: the signaler deems the trial too hard and skips the trial.

utility calculation determines the portion of the environment where each agent is responsible for achieving the target. Here we start from a literal receiver (level-0), on top of this, more complex layers of signaler and receiver are successively built until level-2 reasoning is reached for both agents.

4.8.1 Results

In general, deeper recursive reasoning leads to an increase in performance. For both models, the most complex signaler/receiver pair (level-2 signaler, level-2 receiver) performs best regardless of the environment. This is consistent with previous findings on recursive reasoning where deeper recursion improves performance (Yuan, Monroe, Bai, & Kushman, 2018). When comparing the most complex pair to the simplest (level-1 signaler, level-0 receiver), IW achieves an average of 19.3% and 25.4% boost in performance from recursion in the receiver barrier and signaler barrier conditions respectively. aRSA sees a 20.9% and 15.6% bump in performance. Figure 4.5 shows heatmaps of performance (utility as a percent of optimal) comparing the two models (IW: left, aRSA: right) and two barrier configurations (receiver barrier: top, signaler barrier: bottom) at different levels of recursion. For each pair of signaler and receiver reasoning level (receiver is on x-axis, signaler on y-axis), the utility as a percent from optimal is recorded. A deeper red in the heatmap corresponds to higher performance achieved by a pair of communicators.

Within a signaler level, as the receiver does deeper reasoning, the achieved utility tends to increase (see Figure 4.5). This indicates that having an intelligent receiver is important to performing well on the task. Notably, for both models and environments, the worst performing pair is a level-2 signaler with a level-0 receiver. This could indicate that when the speaker expects their receiver to be reasoning more deeply than they actually are, this mismatch in expectations can be highly detrimental.

At the same level of recursion, IW always outperforms aRSA (see Figure 4.6), achieving up to twice the utility. In fact, the most complex reasoning under aRSA does worse than the simplest IW communicator pair. For IW the simplest reasoning achieves 77.7% (CI: 73.1-82.3%) and 64.0% (CI: 58.8-69.3%) of the optimal achievable utility in the RB and SB conditions respectively. In contrast, the most complex communicator pair under aRSA only reaches 50.0% (CI: 44.7-55.3%) and 44.5% (CI: 39.3-49.7%). This large performance difference indicates that the benefits of recursion are outweighed by the benefits of joint utility reasoning. Here much of the complex inferential burden of language can be pushed to a much simpler utility calculus. If these results align with future empirical behavioral data, it would provide evidence that everyday language does not need deep recursion to be sparse and successful.

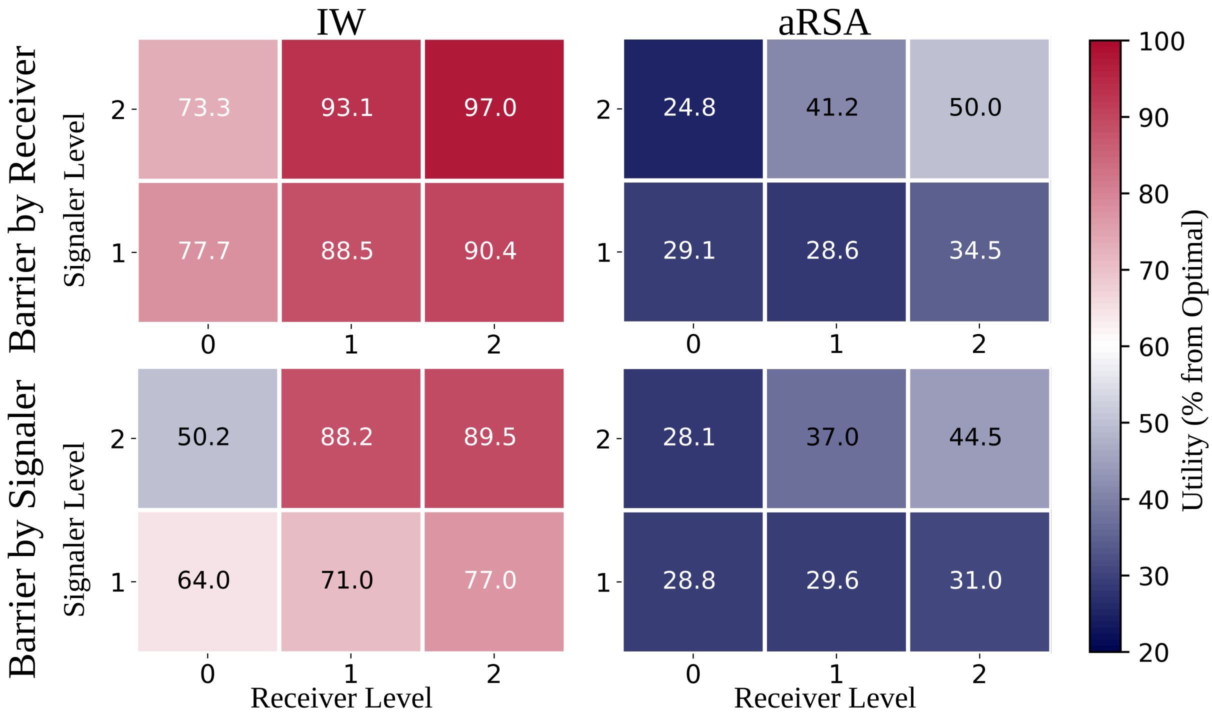


Figure 4.5: Mean achieved utility (red represents higher performance) for signaler and receiver pairs with different levels of reasoning. Shown are N = 500 cases where communication is optimal per modeling level pair. Number of items is fixed at 6, $\beta = 4$.

Finally we can examine the effect of moving the barrier on performance. From a joint utility perspective, moving the barrier toward the receiver makes it harder to constrain the meaning of a signal using joint utility. We find that performance for a communicator pair is better in the RB condition than in the SB condition in IW ($p_{adj} < .001$ for all communicator pairs using Benjamini-Hochberg procedure) but not in aRSA ($p_{adj} > .05$ for all communicator pairs), demonstrating the gains from joint utility reasoning under a shared agency framework.

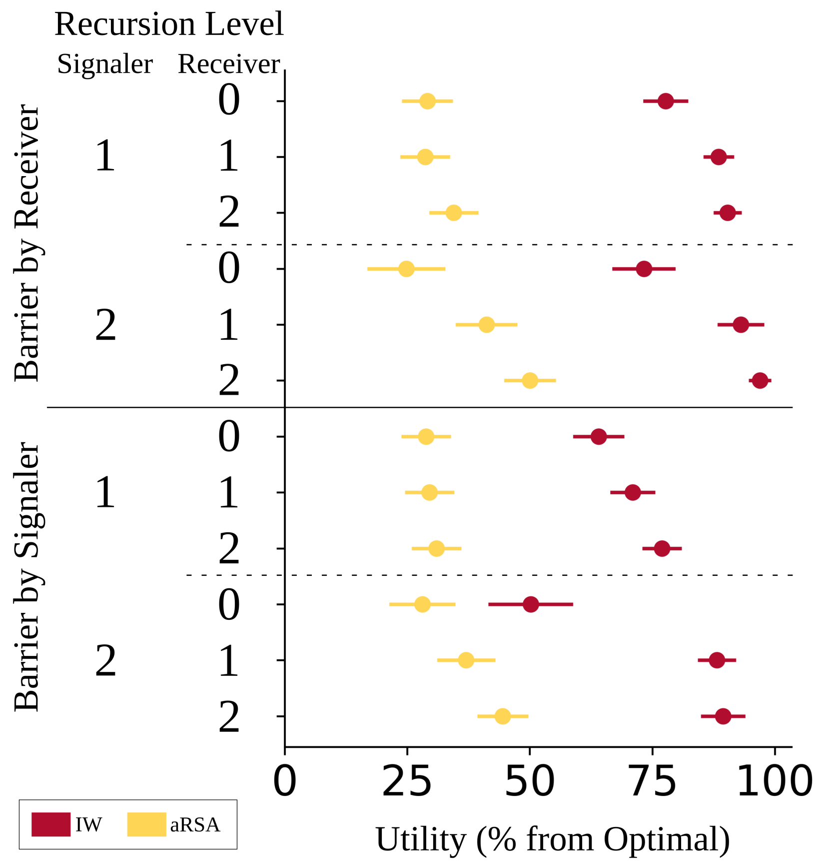


Figure 4.6: Forest plot comparison of IW and aRSA mean achieved utilities (with 95% CIs) at different levels of signaler and receiver recursion and in different barrier positions.

4.9 Simulation 3A: Information Theoretic Gains to IW Under Costly Signaling

Previously, we have demonstrated IW's success in cases where communication is highly restricted. We are now interested in how IW performs when precise, non-overloaded communication is possible but there is an added cost for longer signaling. This still encourages shorter communication as long as it can be interpreted. We predict that communication enabled by the IW will often be shorter because of IW's underlying explicit model of shared agency which can process and integrate context from joint utilities and pragmatics. IW signalers count on their receiver to be intelligently making inferences to understand the signal which removes some of the burden to fully and unambiguously specify what she means off of the signaler.

We aim to show IW is more efficient when compared to a baseline and that this efficiency comes at a minimal cost to task performance. From an information theoretic perspective, we can measure the communication efficiency as the length of communication in bits. We predict IW can communicate using fewer bits while still maintaining task performance.

Previously, each item was defined by values along only two feature dimensions. As a result, even a two word signal is guaranteed to fully specify the target. To motivate signals of varied lengths, we expand the item feature space and make the overloading more challenging. We set both the number of feature dimensions (e.g. color, shape, pattern, size) and the number of values within each dimension (e.g. green, circle, striped, small) to be four. A signaler can say as many dimensions of an item as desired (i.e. up to four words); however, each additional word has a cost. We look at costs from 0 to .6 in increments of .1, where the physical step cost on the grid is still 1.

To make the game challenging and highlight the potential of IW for sparse signaling, we look at environments where, in order to fully specify the item, all features of an item must be signaled in order to uniquely identify the target item. Under these conditions, we compare the average bits used by IW to the full signal. We refer to this fully specified signal as the baseline.

We calculate the average bits of a signal by constructing a predefined codebook given the fact that the distribution of features in the environment is uniform. Thus we would expect the length of each signal to be the same. Each spoken word a signaler communicates corresponds to a combined feature code and value code. As a result, each possible item that could occur in the environment is assigned a unique binary code starting from feature dimension and followed by feature value.[1]

[1] In an environment with n feature dimensions, we encode each possible dimension using a binary representation of a fixed length. For example, a possible mapping might choose 000 to represent shape, 001

As before, all analysis is restricted to cases where communication is optimal, meaning the receiver is in a better position than the signaler to walk to the target item. We allow both the baseline and IW signalers to go to the target item (DIY) and quit when communication is deemed to be too costly or uncertain. Both models compare the utility of signaling to the utility of quitting and DIY. A signaler will go to the target if the utility of signaling (including signal costs) is less than the utility of DIY. In addition, it will choose to quit when the utility of signaling is expected to be less than zero. The only difference is that IW can send shorter signals. Thus we may expect that, when IW can express itself clearly with a shorter signal, it may choose to do so instead of quitting/DIY.

We first investigate what the actual model behavior breakdown looks like, similar to analysis in Simulation 1. Specifically, we are interested in cases where the goal is not reached by the receiver, which are the trials that would be considered sub-optimal from a central control perspective. There are a variety of reasons why the receiver may not reach the goal: the receiver could fail to understand the signal (communication failure), the signaler could decide signaling is too expensive or uncertain and choose to walk to the item (DIY), or the signaler could quit the trial and earn 0 utility (quit). Of these behaviors, communication failure is the most concerning. Moreover, the cases where quitting and DIY do occur may not necessarily be sub-optimal. Although the signaler is physically further from the target, the additional incurred costs of signaling can change that relationship when the cost differential is not large; in some cases, the additional costs of sending a signal can make the achieved utility overall negative or smaller than the signaler doing it herself, in which case it is reasonable to quit and DIY respectively. In addition to model behavior, we compare the average bits needed as well as the overall task performance, measured as utility achieved across different

---

to represent color, 010 to represent size, etc. This allows us to specify which channel we are trying to communicate about. Within each dimension, we then assign each possible feature value a unique code of a fixed length in a similar fashion (e.g. 000 represents the first value, 001 the second, etc.). For example, if the signaler decides to say "red," and red is the first feature value within color, this might correspond to the code '001000.' Alternatively, if the signaler says "red square," the corresponding signal might become '001000 000010'.

signal costs.

4.9.1 Results

Overall, IW successfully communicates in 93.8% of trials whereas for the baseline, successful communication occurs in 88.1% of trials. Figure 4.7 focuses on the cases where the goal is not reached by the receiver at different signal costs. Notably, there are no cases at any cost where the receiver does not understand the signaler's message, indicating that IW will only send a signal when it is confident the receiver will understand. Using IW improves all behaviors. Adjusting for signal cost, IW is significantly more likely to communicate ($OR = 2.18$, $p < .001$)[2], less likely to quit ($OR = .18$, $p < .001$)[3] and less likely to DIY ($OR = .69$, $p < .001$)[4]. In all instances of quitting and DIY observed in IW, sending the full signal would achieve a lower utility than the observed decision. Thus, the signaler only decides to DIY because signaling becomes too costly, or, in rare cases, when the expected utility of signaling as well as acting is negative, the signaler decides to quit.

Notably, the receiver always reaches the target when the cost of each signal is .2 or less. This is because .2 acts as a cost threshold that changes the dynamics between signals and actions: at above a cost of .2, a signal that specifies all feature dimensions becomes greater than one step in the physical environment.

In addition, when we examine only the trials where both models choose to communicate, IW consistently uses fewer bits on average than baseline as shown in Figure 4.8. When there is no cost, saying more will not hurt and there is no incentive for IW to send a shorter overloaded signal, even if it is expected to be understood. As a result, IW only samples shorter signals when they have just as much power to disambiguate as longer ones. When

[2]Model: $logit(Communicate) \sim \beta_0 + \beta_1 model_{IW} + \beta_2 sigCost$

[3]Model: $logit(Quit) \sim \beta_0 + \beta_1 model_{IW} + \beta_2 sigCost$

[4]Model: $logit(DIY) \sim \beta_0 + \beta_1 model_{IW} + \beta_2 sigCost$

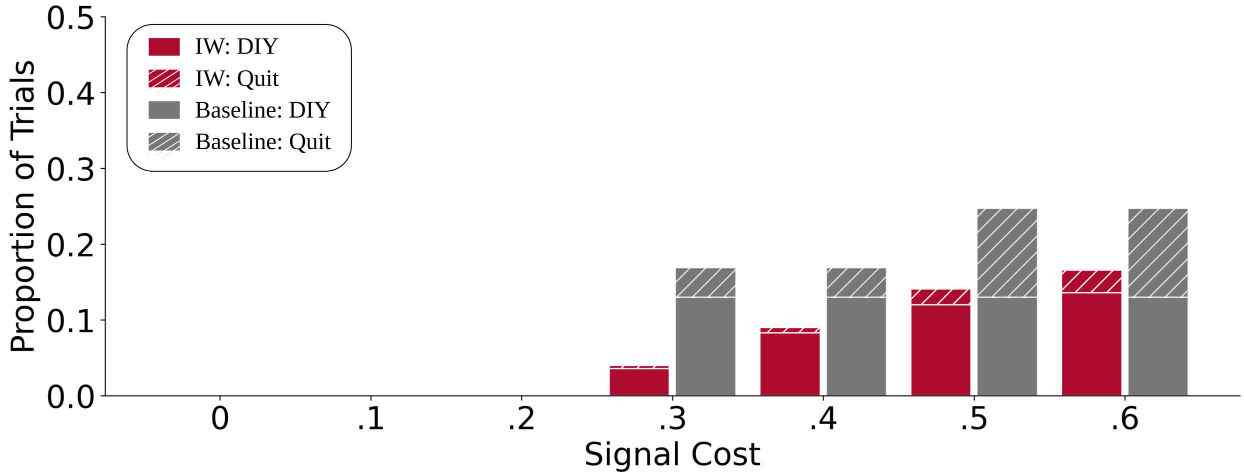


Figure 4.7: Breakdown of cases where IW and Baseline receivers do not go to the target item across signaling costs. Hatched portion of bars represent when the signaler quits and solid portion represents when the signaler chooses DIY.

even a small cost is added, IW further shortens the signals sent. This small added cost serves to significantly amplify the advantage of IW ($p < .001$). As we continue to increase the cost of signaling, however, there is no further reduction of signaling length. This is because with even the smallest added cost, IW is doing the best it can at sending brief signals while still ensuring that they will be understood. This captures the intuition that talking is cheaper than acting, though not completely free. A case where the receiver chooses the wrong target due to signal misunderstanding is much more expensive than sending a longer signal. In fact, sometimes the cost of being uncertain about whether communication will succeed is even larger than DIY or quit, leading the signaler to choose those options.

While IW is more efficient and chooses to signal more often, this could come at the cost of performance in the task. We show this is not the case here. In Figure 4.9 we compare the achieved task utility of IW to the baseline under different signal costs. The total height of each bar represents the action utility achieved as a proportion of the central controller solution, excluding signal costs. It is important to note here that the central controller solution assumes complete, shared knowledge which allows perfect coordination in all trials,

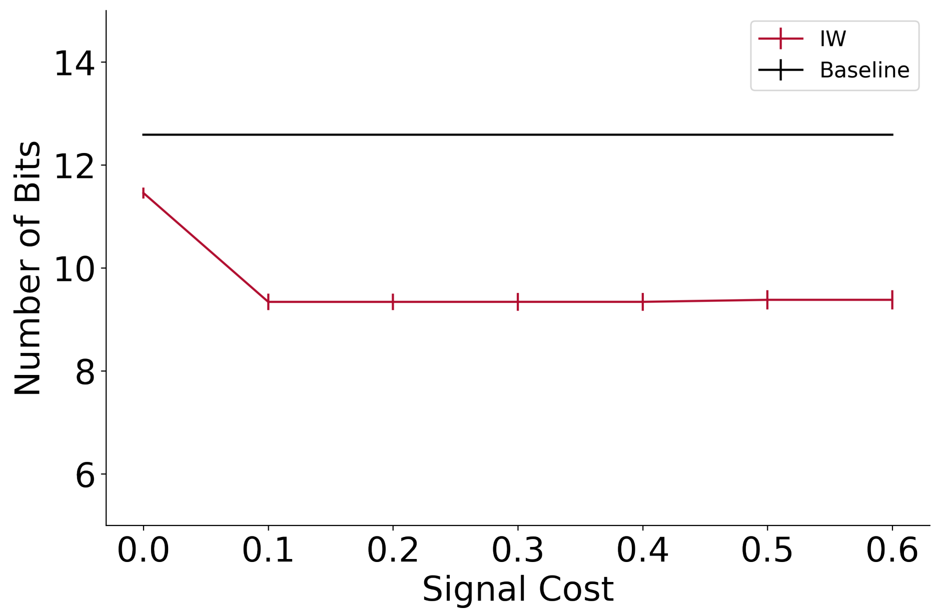


Figure 4.8: Average number of bits used for baseline and IW across different signaling costs (with 95% CI)

without communication. The hatched portion of the bar represents average signal costs incurred. At all signal costs, IW has both a higher total action utility and smaller average signal costs.

In conclusion, through the first portion of this simulation, we have demonstrated that IW can conserve information by sending shorter signals that can be understood despite overloading. As a result, even at large signal costs, IW signalers also DIY and quit less frequently than baseline, leading to increased task performance.

4.10 Simulation 3B: Generalizing the Number of Features

While we have shown the advantages of IW as soon as even a small signaling cost is added, we are also interested in how this effect generalizes across different environments. Here we focus in on this question by manipulating both the number of feature dimensions (e.g. color, shape, pattern, size, edge line weight) and the number of values within each dimension (e.g.

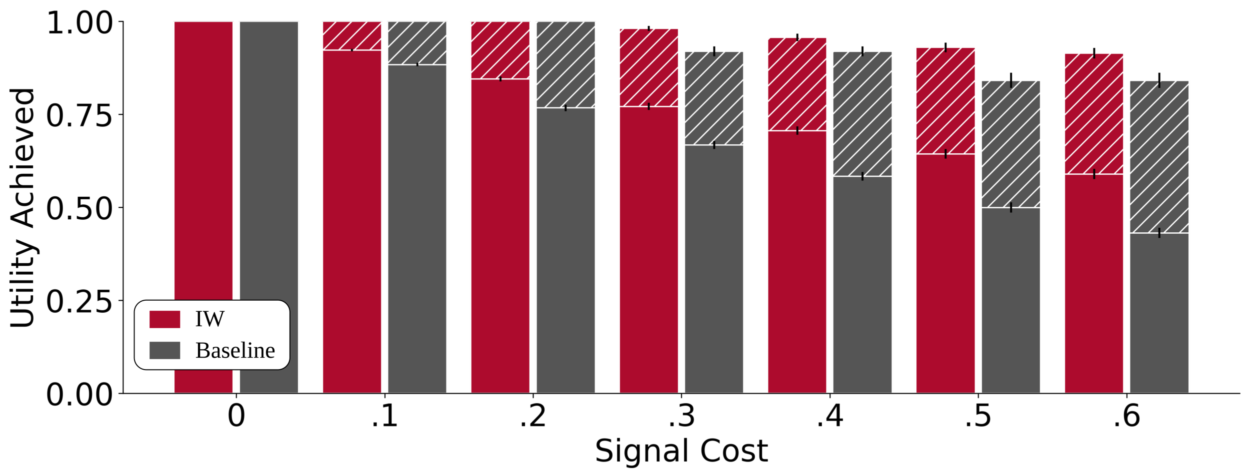


Figure 4.9: Total action utility achieved by IW and Baseline as a proportion of the central controller solution across signal costs. Hatched portions of the bars represent average total cost of signaling incurred at that level of signal cost.

green, circle, striped, small, no edge) from three to five. We fix cost to either be zero or .1 as additional incurred signaling costs do not lead to further reductions in the number of bits conserved. In this focused contrast, we again look at the information conservation and achieved task utility.

4.10.1 Results

We first compare the percent of bits conserved by using IW instead of baseline across different environments for either no signaling cost or a small .1 signaling cost (see Figure 4.10). The resulting heatmap has several important features. First, at both no cost (left) and .1 cost (right), the percent of bits conserved are positive indicating IW is using fewer bits than baseline in all cases. Second, with the small added cost, the percent information conserved is 2.63 times greater on average: this can also be visually observed by how much darker the values in the added cost scenario are. Finally, the amount of information conserved given a signal cost is highly stable. As the feature space increases, IW is still able to send shorter signals that can be understood.

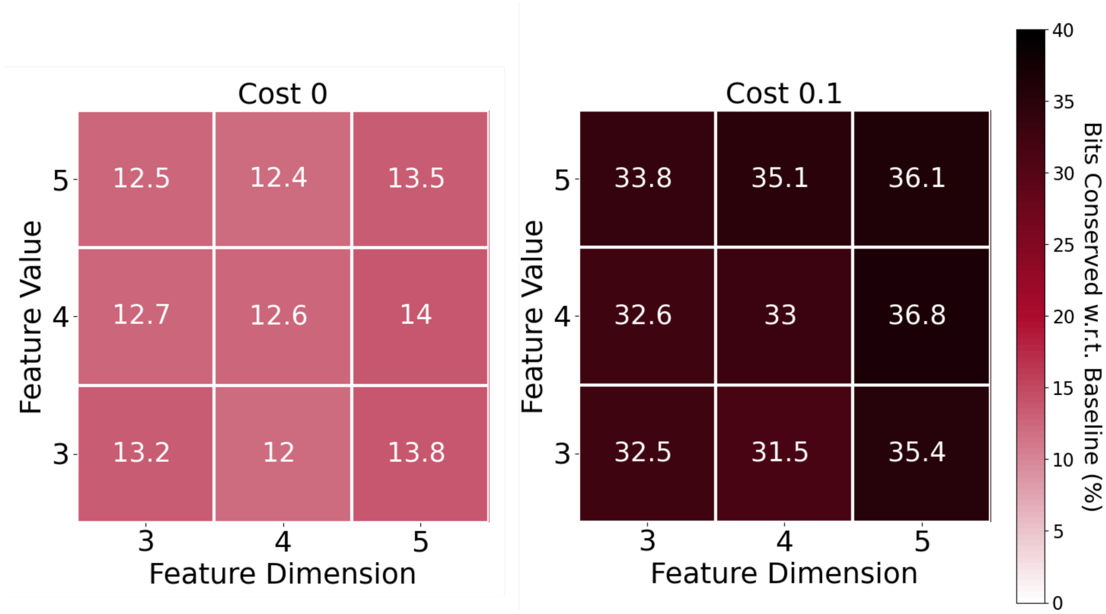


Figure 4.10: Average percent of bits conserved by using IW as compared to baseline for each combination of feature dimension (e.g. color) and feature value (e.g. red). On the left is the percent conserved when the signal cost is zero and on the right is the percent conserved when the cost of each word is 0.1.

While using IW can conserve information, our baseline comparison achieves perfect accuracy at the task which is not guaranteed by IW. Again, we ask whether conserving information comes at the cost of performing well in the task. To answer this, we compare the task utility of IW relative to the baseline as a percent. In this case, an achieved utility of 100% would represent IW is performing on par with the baseline at the task. We see in Figure 4.11 that in all combinations of feature value and dimension IW performs significantly better than baseline (all $p_{adj} < .001$) That is, taking communication costs into account, IW can send shorter signals that are still understood and interpreted by the receiver.

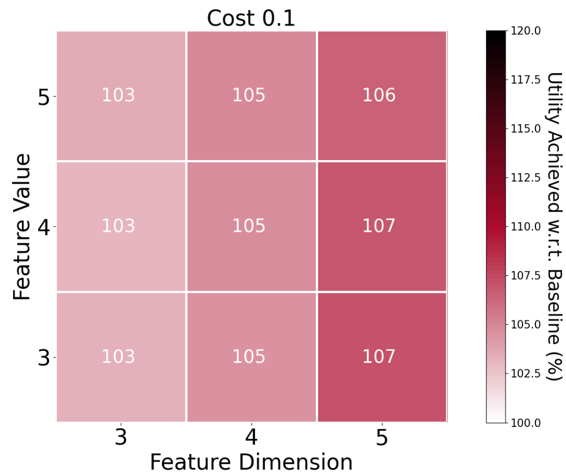


Figure 4.11: Percent task utility achieved by using IW as compared to baseline for each combination of feature dimension (e.g. color) and feature value (e.g. red).

4.11 Discussion

Our proposed modeling approach emphasizes insights from AI and cognitive science to build a model of cooperative communication in grounded tasks based on a shared agency perspective. Cooperative logic, pragmatic language reasoning, and affordable actions under a joint utility calculus constrain a signal's interpretation. Integrating these sources of context allow for fast, flexible signaling which helps remove the inferential burden from deep recursion. As a result, IW serves as a powerful general framework of indirect and overloaded signal production and understanding. We have demonstrated this point through three simulated experiments. Specifically, we showed IW can communicate successfully when highly overloaded, unlike JU and aRSA comparisons, due to its integration of both linguistic and non-linguistic context under a shared agency. Moreover, this consideration of context can take the burden of understanding off of deep recursive reasoning, allowing it to perform well with very little recursion. Finally, when communication is not restricted but incurs a small cost, IW is brief, sending signals using fewer bits while still maintaining high interpretability and task performance.

In the introduction, we highlighted three key properties of IW which allows it to successfully model efficient communication in overloaded scenarios. Our simulated results provide the empirical support for these claims. The first claim was that treating signals as rational actions designed to change others' minds provides a principled framework to send and interpret signals that relies on the assumption of mutual rationality. Reasoning about others' minds allows meaning to flexibly map onto a signal in the context of the situation. The task design allows this by using an environment with action costs (as well as signal costs). Thus, individuals must consider this in order to make rational decisions. This task also places signals in a broader context by integrating them with actions: it is no longer just reasoning about what you want to say but also about what you want to do. This is demonstrated by examining the behaviors different models adopt in Simulation 1. Moreover, the advantage of a rational agentic approach to signaling is observed in Simulation 3 where adding a small signal cost makes communication under IW more efficient and necessarily overloaded without compromising understanding.

Another key feature of IW is that it integrates non-linguistic context about the costs of interacting with the environment into its model of communication. This is motivated by the idea that most communication does not happen in a vacuum – it happens in a shared visual scene full of context. In this task, cooperative reasoning about actions in the environment characterized by joint utility calculations can do much of the heavy lifting in language understanding. In Simulation 1, as the amount of overloading increases, comparisons to JU and aRSA demonstrate that both linguistic and non-linguistic reasoning improve the ability to resolve overloaded signals. The two positions of the barrier in the environment also serve to change the utility dynamic. As a result, IW's advantage is particularly prominent when the barrier is near the receiver where there are more constraints on what the signaler can say under a cooperative joint utility logic. These findings support an account of communication that is able to integrate and process multiple types of relevance for rich understanding despite sparse, indirect signaling.

The final feature of IW is that it adopts a shared We perspective, beyond that of any single individual. This shared agency framework poses extra cooperative constraints on agents, further narrowing the scope of what can be considered rational when producing and interpreting signals. Both IW and JU have a notion of shared-ness through joint planning. However, JU does not involve recursive pragmatic reasoning. Even with a basic model of cooperative logic, JU still achieves limited success and is comparable to aRSA performance in many environmental conditions of Simulation 1. Moreover, individual planning and pragmatic partner reasoning in aRSA are not enough to reliably reduce signaling overloading either. Only when you combine these ideas as in IW do we see robust performance. Finally adopting a shared agency perspective replaces the need for deep recursion, as seen in Simulation 2 where even at very shallow levels of reasoning, IW outperforms a more sophisticated aRSA model. Instead of debating how deep recursion should be, IW shifts the focus to the structure of the minds that are being reasoned about. Thus, by reasoning about a richer shared mind in a broader context, it does not need to go as deep.

CHAPTER 5

Behavioral Experiments: Solving Signaling Ambiguity Through Belief-driven and Action-driven Cooperative Logic

Resolving overloading in communication requires attention to context. Previous research has found that the mutual assumption of cooperation during communication can act as a powerful constraint, allowing successful resolution under ambiguity. In this study, we investigate two specific types of cooperative context used in a communicative task which arise from different sources: beliefs and actions. In pragmatic belief-driven communication, signals are interpreted in context of what else a speaker could have said about the world. Here communicators assume that the speaker aims to change the listener's beliefs by providing the most straightforward signal. In joint utility action-driven communication, signals are considered in terms of what a speaker can reasonably ask others to do given the costs of acting in the physical world. Through a communication game, we tested how listeners would interpret an ambiguous signal using belief pragmatics or joint utility strategies. In Experiment 1, we find that individuals are able to use both strategies and that they are internally consistent about the strategy they choose. Moreover, when these strategies come into conflict, participants are faster and more confident when making joint utility action-driven decisions. Joint utility reasoning is robust across conditions and, in a follow-up which replicates and extends results from Experiment 1, joint utility is shown to be an overall dominant strategy in the population.

5.1 Background

Cooperation has been viewed as a key aspect of communication, providing another form of context to constrain potential ambiguity. Here we collect empirical behavioral data which can lead to insights about how communication operates under this cooperative frame. Moreover, we make the distinction between two types of rational cooperative logic: speech acts and joint planning. Speech acts involve reasoning about signals as a cooperative way to change the *beliefs* of others. Joint planning involves assuming cooperators will choose *actions* that are jointly efficient and fair. These discrete but complementary views offer distinct mechanisms to constrain how signals can be sent and interpreted to resolve ambiguity. While both cooperative aspects of communication have previously been explored, they have been typically viewed separately and from different contexts. In the present study we incorporate them in the same behavioral task to explore whether humans can flexibly employ these two cooperative heuristics for disambiguation based on the context they are in. In addition, when both strategies can be used to solve the task but provide conflicting answers, we examine whether one strategy is dominant.

5.1.1 Context of Beliefs: Cooperative Speech Acts

The first type of cooperative logic employed during communication is speech acts. Speech acts fall under the umbrella of language pragmatics – the branch of linguistics which focuses specifically on the context sensitive interpretation of utterances. Grice's insights in developing a cooperative framework for communication have been highly influential in guiding a formalization of pragmatics. Specifically, Gricean cooperative logic treats communication as a truthful, concise, relevant, and straightforward exchange (Grice, 1975). To be considered cooperative, a signal should be straightforward, maximally efficient, and predicted to be interpretable by the receiver. In order to determine what is straightforward or efficient, communicators must engage in social reasoning about their partners. Although the signaler must ultimately decide on a signal, this process implicitly requires considering the context of all available — but not chosen — options. As a result, a signal with multiple literal meanings may now have a clear pragmatic interpretation which can be inferred using the situational context of the utterance.

While Grice's maxims are intuitively important for communication, alone, they are not enough to solve uncertainty in communication. Instead they must be combined with the insight that exchanges center around the *use* of language. This is valuable because viewing communication through its use ties signals to communicative goals, making their utilities easier to define (Allen & Perrault, 1980; Goodman & Frank, 2016). Under this formulation, communication is a type of rational action: a speech act (Austin, 1962; Clark, 1996; Grice, 1975). When viewed as such, signals have the communicative goal of conveying information about a referent or state of the world to a listener given the decision context (Van Rooy, 2003). A rational, utility driven signaler chooses a signal by evaluating all possible things she could say and picking a good option. Having a communicative goal provides the mechanism for that evaluation of what is good: a signal's value comes from how it is expected to change the listener's beliefs to reflect the intended referent. In turn, under these same assumptions, the listener can use these cooperative constraints to infer the intended pragmatic meaning of the signal.

Empirical evidence also supports a cooperative pragmatic account of communication in adults. Referential language games provide a controlled environment well suited for studying pragmatic reasoning (Lewis, 1969; Wittgenstein, 1953). In these games, a set of items with different features (e.g. shape, color) act as context, and a listener aims to understand which referent a speaker is indicating from a potentially ambiguous signal. In one game, listeners were asked to bet on which item they believed the signal referred to by distributing money across the different possibilities (Frank & Goodman, 2012). The listeners' bets (combined with empirical ratings of feature salience) agreed highly with how informative the speaker's signal was for disambiguating the item. These effects were also replicated in a forced choice task with a similar setup (Qing & Franke, 2015) as well as in a setting with more complex stimuli depicting ambiguous spatial relations, albeit with more noise (Carstensen, Kon, & Regier, 2014).

5.1.2 Context of Actions: Cooperative Joint Planning

The second type of cooperation we focus on is the context that joint planning provides in a shared task. Much of communication occurs face-to-face where perceptual cues in the environment provide important context for framing an exchange. From this perspective, communication is simply a social tool which can enable individuals to coordinate and get things done together more effectively (Bruner, 1985; Tomasello, 2000; Vygotsky, 1978). Again, communication is framed in terms of use, but this time studied using commonsense knowledge outside of language. Instead, this knowledge lies in considering consequences in the physical world through action planning and in others' mental world which provide the beliefs and desires to create a plan.

We motivate our emphasis on joint action context by examining how even young children who do not yet have the capacity for fully-developed language can intelligently and flexibly reason using sparse communication but using constraints from acting cooperatively, namely through joint commitment and fairness. Before they have mastered language, toddlers can use visual communication to monitor and regulate their partner through protesting or attempting to re-engage them when they break a joint commitment formed through verbal acknowledgement (Gräfenhain, Behne, Carpenter, & Tomasello, 2009; Warneken et al., 2006). At as young as four years old, children already exhibit sensitivity to minimal communication: they establish commitment using simple cues such as joint attention to help offset risks of cooperating in a stag hunting paradigm (Wyman, Rakoczy, & Tomasello, 2013). Moreover, slightly older children protest when their partner does not cooperate, even when eye contact was the only established form of joint commitment (Siposova et al., 2018). These findings indicate that communication can help establish strong joint goals in the context of cooperation.

One of the early non-verbal uses of communication is also demonstrated in the context of fairness. Children, but not chimpanzees are able to split rewards fairly in a collaborative task where it is easy for one party to monopolize rewards (Hamann et al., 2011; Warneken et al., 2011). In the few cases where one child tries to take more than is fair, sparse communication (e.g. "Hey!") quickly and decisively resolves disputes. Here "Hey" is overloaded, and this overloading is not solved by considering alternative protests or signals as in pragmatic reasoning. Instead, it is solved by considering the context of the task — where the principle of fairness is being violated. This protest comes from not only a preference for equality but also a resentment at being treated unfairly (Engelmann & Tomasello, 2019). These developmental studies demonstrate the importance of task-based cooperation in communication stripped down to its most fundamental form, without syntax or grammar.

These cooperative properties of commitment and fairness can be realized through utility driven joint planning: cooperators act under a rational plan that apportions fair costs and rewards to all parties given a joint goal. Empirical evidence has also shown that adults engage in joint utility planning for cooperative tasks, preferring co-efficient actions which prioritized the group utility over the utility of any individual (Török, Pomiechowska, Csibra, & Sebanz, 2019). From a utility driven standpoint, even toddlers understand the cooperative logic of ambiguous requests from a joint cooperative perspective (Grosse et al., 2010). In this experiment, two equivalent items are equidistant from the toddler, but near and far relative to the speaker. When the speaker makes an ambiguous request for the item, children are able to use cooperative logic to reason over the *joint* utility dynamics of the environment in the context of the speaker's capabilities: reaching the far item more often when the speaker had their hands free than occupied. These studies support how communication should be taken in context of committing to achieve a shared goal fairly and respectfully. In both children and adults, joint planning ultimately makes it irrational to ask a collaborator to do something more easily accomplished by oneself.

5.2 Experiment 1: Belief-Pragmatics and Action-Utility Use for Signal Disambiguation

5.2.1 Methods

This task combined feature overloading enriched by a spatial scene, which included abilities to disambiguate signals both using the belief-driven context of words and the action-driven context of utility dynamics. In a grid-world environment, participants played a referential communication game where they were told the goal was to cooperate with their partner to reach a target item in the fewest steps. During the game, the participant always played the role of a receiver who could observe the entire environment but did not know which item was the intended target. Participants were told they were working with a cooperative, intelligent signaler who had a full view of the grid and knew the target; however, in reality, signals were pre-programmed. The signaler's decision depended on the condition and consisted of either an ambiguous signal — consistent with multiple potential items in the trial — or walking to an item when that item was closer to the signaler than the receiver.

5.2.2 Participants

Sixty-six undergraduate students in the Department of Communication at University of California, Los Angeles (UCLA) participated in this online study for class credit. We analyzed the data of 51 participants after excluding six participants for not finishing the experimental trials, three participants for failing the comprehension quiz more than twice, and seven participants for self-reporting not being serious in the experiment (one participant among them also did not finish the experimental trials). The experiment was performed in accordance with guidelines and regulations approved by the UCLA institutional review board IRB#19-001990.

5.2.3 Stimuli and Task

Participants were able to access the experiment on their personal computer or laptop. On each trial, a 9 by 10 grid layout was presented to participants. Each grid square was 50 px $\times$ 50 px and, three items were placed in the grid. Each item had two features of color (orange, purple, or green) and shape (triangle, circle, or square) for a total of nine distinct items. An icon representing the participant was located at grid location (4, 6) while their partner was located at (4, 0). Both agents traveled along the grid taking steps in the four cardinal directions, so Manhattan distance was used to describe how far each item was from an agent.

5.2.3.1 Design

The experiment followed a within-subject design with four conditions: Belief-Pragmatic, Action-Utility, Conflict, and Signaler-walk. Participants played a total of 80 randomly ordered trials (20 per condition). The main dependent variable was the strategy the participant employed to solve each condition, reflected by the item they chose as the target. The receiver's decision time from when they received a signal to when they selected an item was recorded. In addition, participants rated their own confidence after each decision.

The Belief-Pragmatic condition coincided with the example from Frank and Goodman (2012), but was spread spatially in a visual display. All utility dynamics were fixed so that only the features of the items could influence the receiver's decision, equivalent to reasoning in a referential signaling game. Two items had one unique feature and one feature shared with a third item (see Figure 5.1 for an example trial of each condition). The signal was a shared feature, consistent with two items. Relevant items were equidistant from the receiver and all items were jointly efficient for the receiver (closer to receiver than signaler). Receivers could select an item that was irrational: inconsistent with the signal, non-pragmatic: consistent but could be indicated with a more straightforward signal, or pragmatic: consistent and most straightforward because both features were overloaded. In a traditional referential signaling

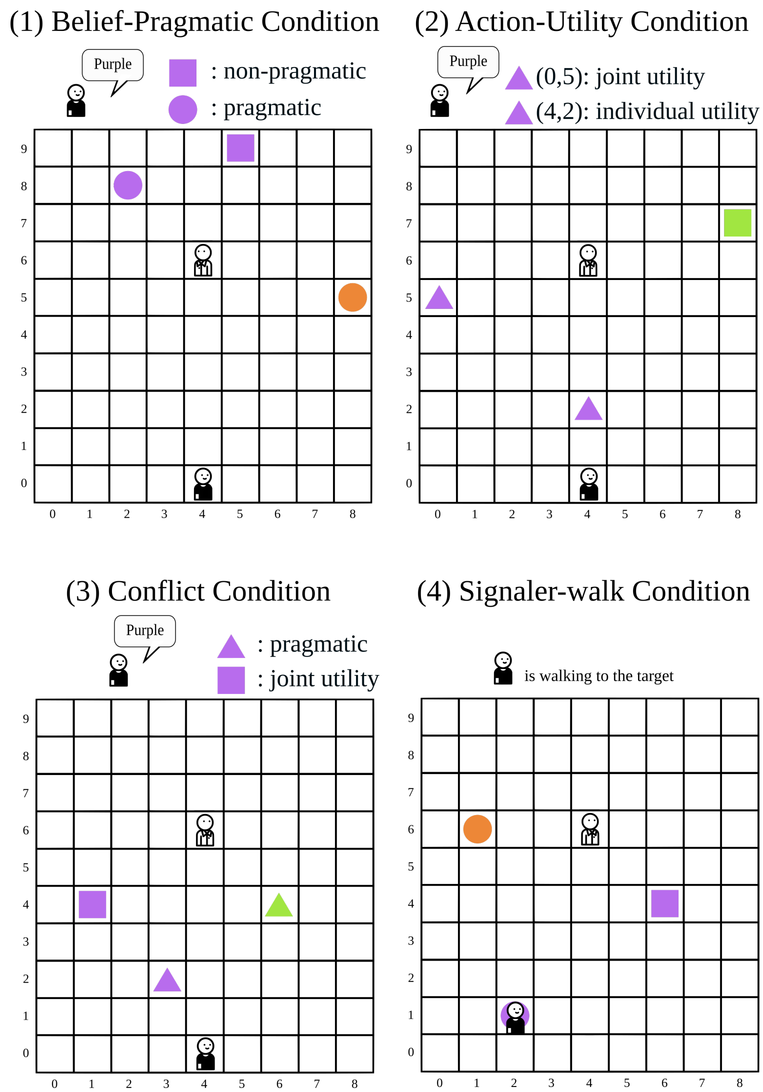


Figure 5.1: Example trials of four experimental conditions. Signaler decision (visible to participant) and reasoning corresponding to item selection (not visible to participant) on top of grid layout. (1) Belief-Pragmatic Condition. (2) Action-Utility Condition. (3) Conflict Condition. (4) Signaler-walk Condition.

game setup, a distinction is made between a literal and pragmatic reasoner whereas we divide responses into pragmatic and non-pragmatic items. This is because selecting either consistent item on an individual trial is consistent with literal reasoning; understanding whether an individual is truly reasoning pragmatically is only possible when looking at the distribution of an individual's response across many trials.

The Action-Utility condition forced participants to make a purely utility-based decision with two identical items (and one irrelevant distinct one). The observed signal was one feature of the identical items, which made the context of language pragmatics unable to help with disambiguation. This setup reflected the dynamic in Grosse et al. (2010), but with a stronger individual utility component. One of the identical items was closer than the other to the receiver, making it individually efficient to reach. However, the individually efficient item was also closer to the signaler than to the receiver, making it jointly inefficient. Thus, receivers could select an item that was irrational: the non-identical inconsistent one; individual: individually efficient but jointly inefficient; or joint: jointly efficient but individually inefficient.

The Conflict condition was designed to force participants to choose between a joint utility and pragmatic strategy. It was identical to the Belief-Pragmatic condition in terms of item feature structure and signal. Also, the two consistent items were equidistant from the receiver. However, instead of all items being jointly efficient for the receiver, the pragmatic item was jointly efficient for the signaler. Receivers could still select an irrational item, but now had two previous strategies that came into conflict and could select either a pragmatic interpretation inconsistent with joint utility (pragmatic) or a joint utility interpretation that was not pragmatic (joint).

Finally, in the Signaler-walk condition, the signaler walked to an item, and participants did not make a decision. This baseline was to establish that the signaler had the option to act for herself instead of communicating. Moreover, the signaler only walked to items when it was jointly efficient to do so, demonstrating the signaler was rational and cooperative.

Items and signals were counterbalanced to account for preference of feature or feature value. In addition, items were separated by a minimum distance of two grid units to reduce potential perceptual chunking. Items always were always at least two grid units farther from one agent than the other in order to ensure clear joint utility judgments. Finally, item locations within a condition were sampled randomly without replacement, subject to the utility constraints defined by the condition and aforementioned restrictions.

5.2.3.2 Procedure

Participants entered the experiment by opening the link on their own device. They started with an instruction tutorial which established the rules and cooperative context of the task, and then completed a comprehension quiz that tested them on the goal and set-up of the experiment. Participants completed eight practice trials to familiarize them with the task which consisted of two trials in each condition presented in a random order.

In each trial, the signaler made the first decision: she either walked to an item herself or sent a signal to the participant describing a single item feature (e.g. "circle"). If the signaler sent a signal, the participant then had a chance to walk to the item they believed was the target by clicking on it. Before they made a decision, hovering the cursor over any item in the grid displayed the distance of each agent from the target: the cost of traveling there. If the signaler moved to the target herself, participants observed the signaler walking to the item. The trial ended when either agent reached an item. Then, a review box would pop up, showing who took how many steps to reach the selected item. Participants were asked to rate their confidence in their selection from one (least confident) to five (most confident). Participants then proceeded to the next trial. Participants were not given feedback about whether their decision was correct or not to avoid biasing decisions on future trials. After all experimental trials, participants took an exit survey which included a self-report on how serious they were throughout the experiment, strategies they used, and performance of their partner.

5.2.4 Results

We analyzed the strategy, response time, and confidence rating on each trial. Across all conditions and participants, only 17 trials had irrational responses (Belief-Pragmatic: 9, Action-Utility: 3, Conflict: 5), thus we restricted our analyses to focus on the major strategies employed: for the Belief-Pragmatic condition, pragmatic/non-pragmatic; for the Action-Utility condition, joint/individual; and for the Conflict condition, joint/pragmatic.

5.2.4.1 Strategy Preferences

We examined the breakdown of each participant's behavior within each condition. In the Belief-Pragmatic Condition, this was the proportion of trials each individual chose the pragmatic versus non-pragmatic item shown in Figure 5.2. If an individual was reasoning literally instead of pragmatically, we would expect them to have no preference between the two consistent items: pragmatic and non-pragmatic. After adjusting for multiple comparisons using the Benjamini-Hochberg criteria (BHC), we find that 20 out of 51 participants (39.2%)have a strong strategy preference ($p_{adj} < .05$). Nineteen of these participants are using pragmatics.[1]

For the other two conditions, the items themselves corresponded to two types of strategies; thus, we tested strategy consistency within an individual in the same manner as well as whether a clear preference emerged across the population. At the individual level in the Action-Utility conditions, 45 out of 51 participants (88.2%) had an individual preference for a strategy (BHC adjusted $p_{adj} < .05$) which can be seen in Figure 5.3. At the same time, people were divided between an individual utility and joint utility strategy. Twenty-four participants (47%) were driven by individual utility and 21 participants (71%) selected items that was consistent with maximizing the joint utility. At the population level, we looked at the distribution of proportions of joint utility use, averaged by subject and performed a

[1]One participant had a strong preference for the non-pragmatic item. This strategy requires performing pragmatic reasoning but deliberately selecting the non-pragmatic option.

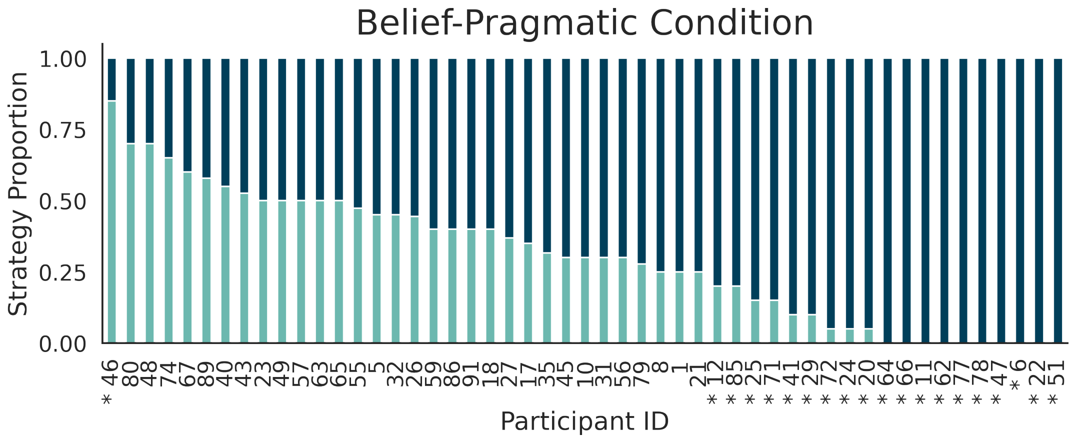


Figure 5.2: Individual participant strategies in the Belief-Pragmatic condition. Asterisks denote participants who were pragmatic reasoners.

t-test under the hypothesis $H_0 : \mu = .5$. Under this test, there was no evidence to support a dominant utility strategy in the population ($\overline{x}_{ju} = .559$, $p = .331$).

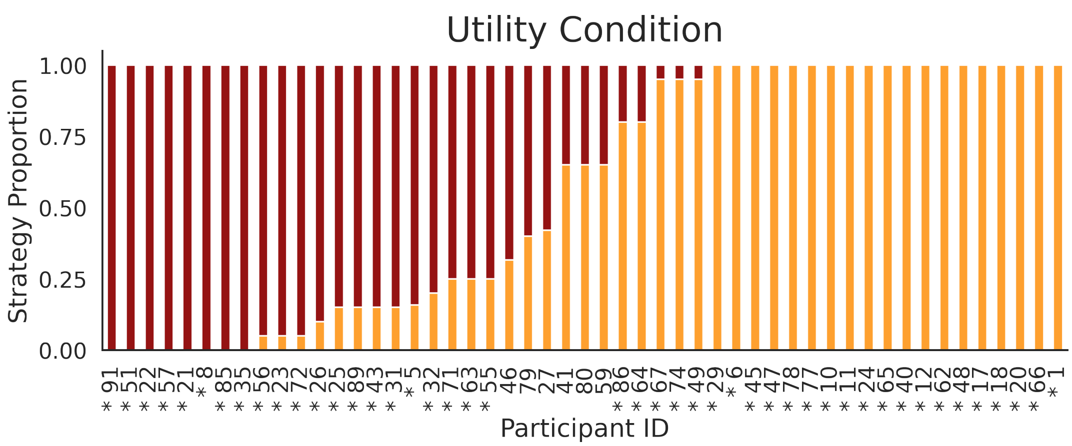


Figure 5.3: Individual participant strategies in the Action-Utility condition. Asterisks denote participants who were highly internally consistent.

Finally, we focus on the results of the Conflict Condition, our main experimental manipulation of interest. At the individual level, people were highly internally consistent when choosing a strategy (see Figure 5.4). In 39 out of 51 cases (76.5%), participants adopted a dominant strategy (BHC adjusted $p_{adj} < .05$). Of these, 23 individuals (45%) used joint utility reasoning while 16 individuals (31%) selected pragmatic items. While there were slightly more individuals who employed joint utility reasoning, there was not enough evidence at the population level to suggest an overall preference ($\overline{x}_{ju} = .569$, $p = .234$).

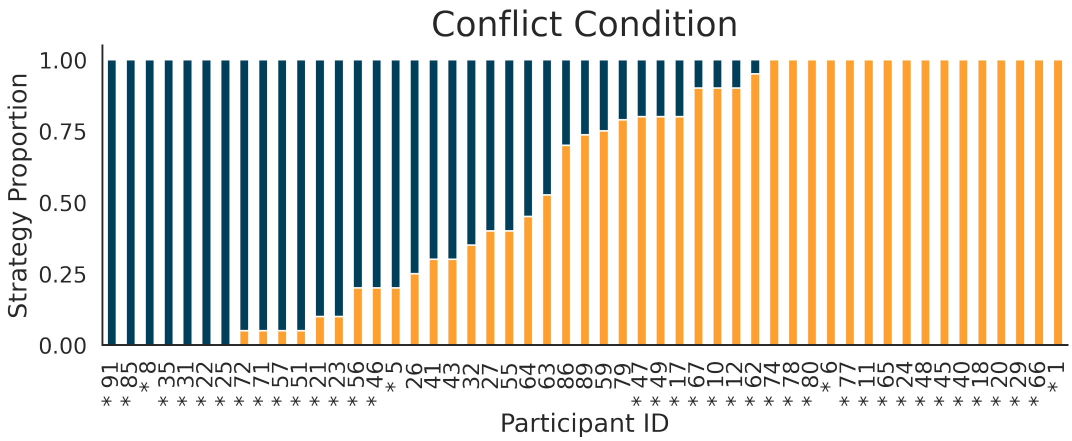


Figure 5.4: Individual participant strategies in the Conflict condition. Asterisks denote participants who were highly internally consistent.

Moreover, we investigated whether participants' strategies correlated between conditions in Figure 5.5. Pairwise correlation analyses indicated a strong positive relationship between an individual's strategy in the Action-Utility and Conflict conditions (Spearman's $\rho = .91$, $p < .001$). That is, individuals who chose a joint utility strategy in the Action-Utility condition were also likely to choose a joint utility strategy when pragmatic reasoning and utility reasoning were in conflict. This effect was not observed for the Belief-Pragmatic and Conflict ($\rho = .03$, $p = .852$) or Belief-Pragmatic and Action-Utility conditions ($\rho = .18$, $p = .211$).

5.2.4.2 Strategy Difficulty: Decision Time and Confidence

In this task, we examined decision time which can act as a rough proxy for the cognitive difficulty involved in employing that strategy (Townsend, 1992). Because participants took the study on their personal device instead of a controlled laboratory setting, the data included

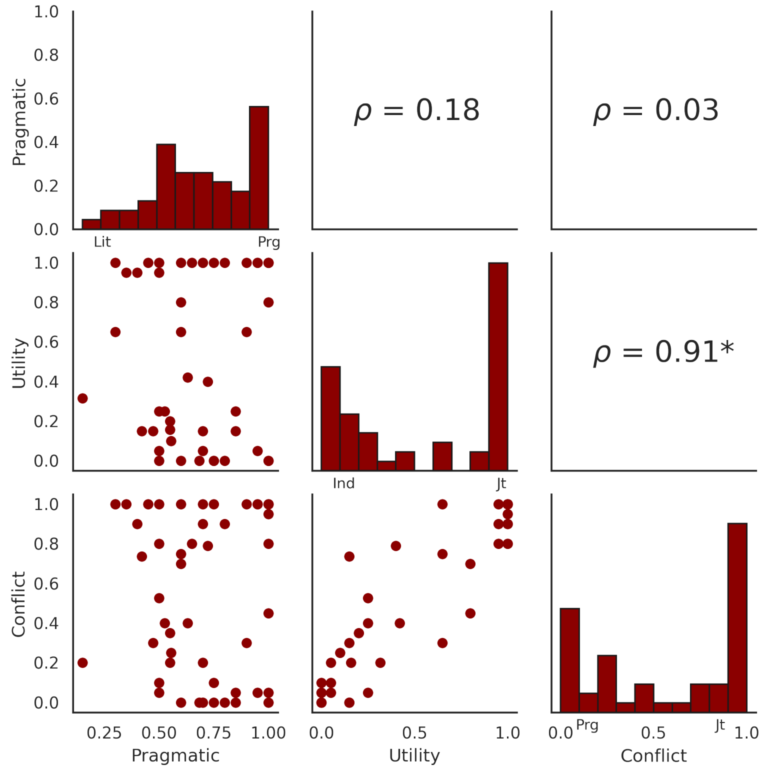


Figure 5.5: Strategy correlations across conditions. Correlation coefficients (upper triangle), corresponding to the individual responses (lower triangle). Histogram describing distribution of strategy preference (on diagonal). Data concentrated at the extremes of the histogram indicate the strong, divergent preferences seen in the Action-Utility and Conflict conditions.

extreme decision times which could not reasonably be attributed to deliberation on the task. While it is common practice to remove reaction times above a certain threshold, typically three Z-scores away (Tabachnick, Fidell, & Ullman, 2007), this can substantively inflate the Type I error rate (Bakker & Wicherts, 2014). Because we had no strong literature-based intuition for a decision time cut-off to indicate when subjects were no longer paying attention, we relied on nonparametric testing which is robust to outliers and skew inherent in reaction time data. In the Belief-Pragmatic condition, we found participants to take more time when employing pragmatic reasoning than non-pragmatic reasoning ($\tilde{x}_{prag} = 5.23$ sec, $\tilde{x}_{lit} = 4.58$ sec, Mann-Whitney-Wilcoxon test (MWW); 95% CI of median difference: [.201, 1.103], $p < .001$). In the Action-Utility condition, participants took similar time to respond when employing either strategy ($\tilde{x}_{ju} = 2.93$ sec, $\tilde{x}_{iu} = 2.83$ sec, MWW; 95% CI: [-0.094, .389], $p = 0.115$). Finally, in the Conflict condition, participants spent longer to make a decision when employing pragmatics as opposed to joint utility reasoning ($\tilde{x}_{prag} = 3.61$ sec, $\tilde{x}_{ju} = 3.07$ sec, MWW; 95% CI: [.003, .561], $p = 0.024$). See Figure 5.6 for decision time comparisons within each condition.

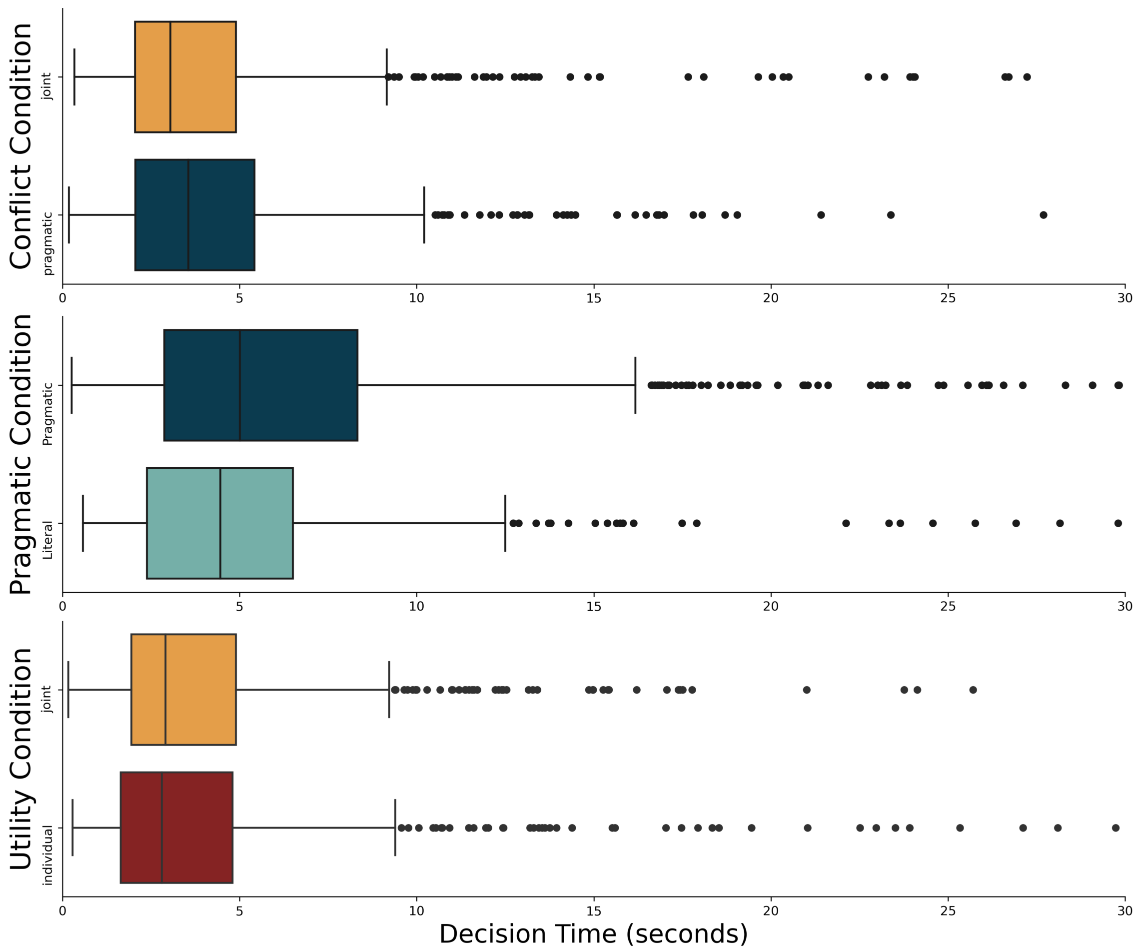


Figure 5.6: Boxplot of decision times for three conditions. Trials > 30 seconds ($n_{pragmatic} = 17$, $n_{utility} = 7$, $n_{conflict} = 9$) are included in analyses but not shown here for legibility.

We also examined self-reported confidence as a function of decision (see Figure 5.7). Participants were significantly more confident when choosing pragmatic items than non-pragmatic ones ($\overline{x}_{prag} = 3.60$, $\overline{x}_{lit} = 3.05$, $p < .001$ under Welch's t-test) as well as when choosing items that maximized joint utility rather than individual utility ($\overline{x}_{ju} = 4.06$, $\overline{x}_{iu} = 3.44$, $p < .001$). Finally, in the Conflict condition, participants were significantly more confident when choosing the joint utility items over pragmatic ones ($\overline{x}_{ju} = 3.99$, $\overline{x}_{prag} = 3.67$, $p < .001$).

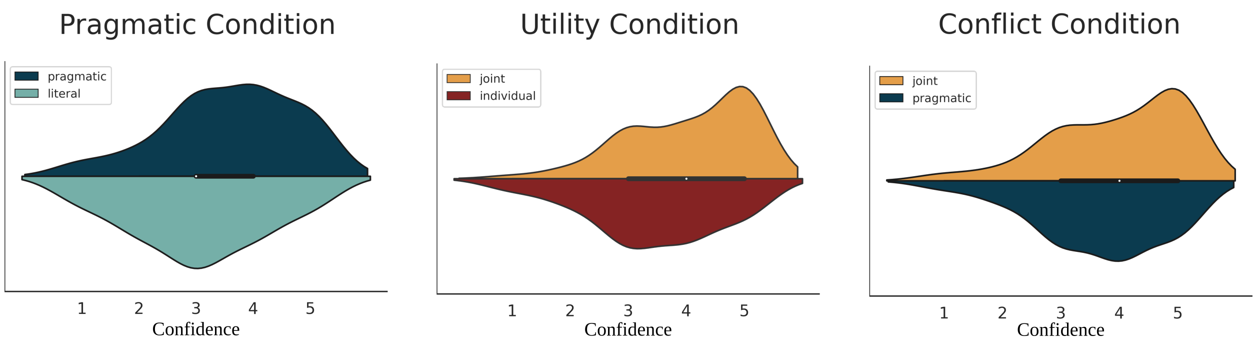


Figure 5.7: Self-rated confidence for each condition on a scale of 1 (not at all confident) to 5 (highly confident). Violin plots are split by strategy to show the difference in confidence distribution depending on chosen strategy.

5.2.5 Debriefing Data: Collaboration Between Partners

While participants were told they were playing with a cooperative and intelligent speaker, without live interaction, there was no way to regulate one's partner. As a result, one potential concern was that participants would not feel that their partner was collaborative or that they may lack the motivation to be collaborative themselves. However, qualitative examination of post-experiment debriefing data suggests participants did engage collaboratively with their partner.

When asked to rate their partner's performance in the experiment on a scale of 1 (lowest) to 5 (highest), participants tended to rate their partners favorably or above average ($\overline{x} = 3.79$). Upon examining participants' written reflections, 36 of the 51 participants (70.58%) indicated that their strategy involved reasoning about their partner collaboratively, beyond just the non-pragmatic signal that was sent. Self-reported strategies included references to joint utility by considering relative distances between their partner and the targets, as well as references to pragmatics by considering what signals their partner did not choose to send in addition to the one that was sent. Finally, when given a chance to report any issues or confusion with the experiment, only two of the 51 participants (3.92%) expressed any doubt over the helpfulness of their partner.

5.2.6 Discussion

Individuals do use pragmatic reasoning, but it is challenging and not universal in this task, which we established through the strategy preference analysis in the Belief-Pragmatic Condition. This supports previous empirical findings in referential signaling games (Frank & Goodman, 2012; Qing & Franke, 2015), replicating this phenomena in our visual paradigm. On average, people also took longer to make a pragmatic decision than a non-pragmatic one, which is consistent with the computational models of pragmatics. In order to come up with a pragmatic interpretation of a signal, a listener must first reason over simpler interpretations (Goodman & Frank, 2016). At the same time, people were more confident about pragmatic selections than non-pragmatic ones.

Finally, while there was not enough evidence to say joint utility was a dominant strategy over pragmatics in this task, when we consider the Conflict condition at the individual level, we see that people are exceptionally strategic in their decisions. Results suggest groups of highly consistent decision-makers who have overwhelming preferences for their respective strategies. Some individuals interpreted signals in a belief-driven manner: reasoning based on the speaker's intention to be straightforward. Other individuals interpreted signals in an action-driven manner: interpreting signals in a way that led to jointly efficient actions.

However, there is further evidence that joint utility may be a more robust strategy than pragmatics. When examining an individual's decisions across different conditions, their use of joint utility in the Action-Utility condition was highly predictive of their behavior in the Conflict condition. In contrast, an individual's ability to reason pragmatically in the Belief-Pragmatic condition had no bearing on their decisions in the Conflict condition. This suggests that while only a subset of individuals used a joint utility based strategy, it was an incredibly powerful and robust heuristic that could generalize across contexts.

Moreover, in the Conflict condition participants were in fact faster at making joint utility based decisions than pragmatic ones. At the same time, confidence ratings were higher on trials where people employed joint utility. These results suggest that joint utility reasoning may be an easier means to solve this task than pragmatics.

5.3 Experiment 2: Focus on Joint Utility and Pragmatics in Conflict

Results from Experiment 1 indicated that when joint utility and pragmatics were in conflict, individuals were faster and more confident about joint utility reasoning decisions. Moreover, while there was not enough evidence for joint utility as a dominant strategy in the population, it tended to be more robust. In Experiment 2 we focus on the contrast between pragmatics and joint utility to replicate and extend existing results in Experiment 1. We also hypothesized that the inclusion of the non-conflict conditions could be explicitly cuing people to consider a strategy that they might not have naturally used otherwise, introducing additional bias.

5.3.1 Methods

In order to focus on the conflict between pragmatics and joint utility and remove a potential source of bias, we removed the Belief-Pragmatics and Action-Utility conditions from the experimental setup and just focused on cases where subjects were forced to make a conflicting choice (Conflict condition). We also kept the Signaler-walk condition to demonstrate that the signaler could act cooperatively. Other than this change, the stimuli, task, and game play were all the same as in Experiment 1. As a result, a new set of participants saw 20 trials of each of the two remaining conditions for a total of 40 trials.

Fifty undergraduate students in the Department of Communication at University of California, Los Angeles (UCLA) participated in this online study for class credit. We analyzed the data of 40 participants after excluding three participants for failing the comprehension quiz more than twice, and seven participants for self-reporting not being serious in the experiment.

5.3.2 Results

We performed the same analyses as in Experiment 1 with particular emphasis on the overall strategy and individual strategy preferences. Across all conditions, four trials had irrational responses which were excluded from analysis. As before, individuals still tended to be consistent in choosing their strategy (see Figure 5.8). In this case, 26 out of 40 (65%) participants adopted a dominant strategy after adjusting for multiple comparisons using BHC ($p_{adj} < .05$). Of these, eight (20%) adopted a pragmatic strategy and 18 (45%) adopted a joint utility strategy. In this case, unlike the previous experiment, a two-tailed t-test of subject strategies indicated the joint utility was dominant ($\overline{x}_{ju} = .623$, $p = .0397$).

Furthermore, we replicated the decision time and confidence patterns seen in Experiment 1. Participants spent longer to make a decision when employing pragmatics as opposed to joint utility reasoning ($\tilde{x}_{prag} = 4.20$ sec, $\tilde{x}_{ju} = 3.58$ sec, MWW; 95% CI of median difference [.069, .857], $p < .001$). Additionally, they were more confident when choosing the joint utility items than when choosing the pragmatic ones ($\overline{x}_{ju} = 4.03$, $\overline{x}_{prag} = 3.70$, $p < .001$ under Welch's t-test).

Finally, participant ratings of their partner during the exit survey in Experiment 2 matched those of Experiment 1. When asked to rate their partner's performance in the experiment on a scale of 1 (lowest) to 5 (highest), participants tended to rate their partners

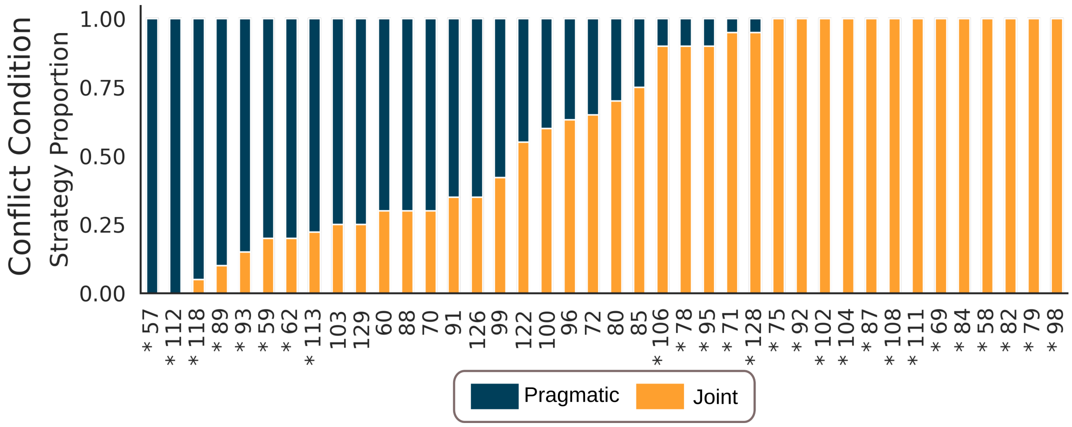


Figure 5.8: Strategy in Conflict Condition: Individuals are highly internally consistent and have an overall preference for joint utility. Asterisks denote participants who were highly internally consistent.

favorably ($\overline{x} = 4.10$). Thirty of the 40 participants (75.0%) indicated that their strategy involved reasoning about their partner collaboratively, with similar strategies to those reported in Experiment 1. No participants reported doubt over their partner's cooperation.

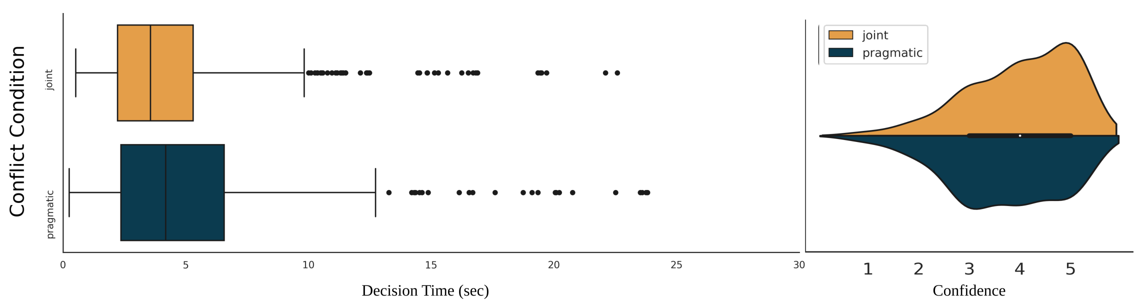


Figure 5.9: Left: Boxplot of decision times. Trials > 30 seconds ($n = 9$) are included in analyses but not shown here for legibility. Right: Self-rated confidence, split by strategy.

5.3.3 Discussion

These results serve to both replicate previously described patterns observed in Experiment 1 as well as extend them. After removing a source of potential bias generated from the inclusion of trials testing for utility reasoning and pragmatic reasoning individually, there are multiple measures of evidence which all support joint utility as a dominant strategy when compared to pragmatics. First, all results in this follow-up remain consistent with the original experiment. Around two-thirds of individuals are highly individually consistent in their strategy, and participants who adopted a joint utility strategy were both highly confident and faster to make their decisions. Additionally, self-reported strategies again affirm that the majority of participants believe their partner to be at least somewhat collaborative and engage in decision-making centered around this belief. Finally, in addition to the replications from Experiment 1, we also see a preference for joint utility reasoning at the population level.

5.4 General Discussion

When individuals were forced to choose between action-driven or belief-driven strategies in Experiment 1, they were highly internally consistent. Moreover they were faster and more confident about their action-driven strategies using joint utility reasoning. In Experiment 2 we were able to confirm these results as well as extend them. When we focused in on this Conflict scenario, joint utility began to dominate as a strategy.

Individuals' overall preference for joint utility reasoning was not as strong as initially hypothesized. For example in the Action-Utility condition of Experiment 1, we found participants to be more confident about their decision when they chose a jointly efficient item, but that they were still highly split between an individual and joint. Moreover, in the Conflict conditions, it was not until Experiment 2 where we observed a population level effect. We were surprised that this joint utility preference was not stronger given previous work on joint efficiency in cooperative tasks (Török et al., 2019). One explanation is that being cooperative requires effort. In this experiment utility decisions can be mapped on to different types of perspectives. Maximizing individual utility is equivalent to viewing the problem from an egocentric perspective. There is a line of research that suggests reasoning from an egocentric perspective acts as an easier default which only adjusts toward another perspective with sufficient cognitive resources (Keysar, Barr, & Horton, 1998; Epley, Keysar, Van Boven, & Gilovich, 2004; Barr, 2014). In this task, it is possible that participants did what was easiest. However, this may not be the whole story as (a), there was no significant difference in the distribution of decision times between individual and joint utility, and (b), participants were actually more confident about joint utility decisions. These findings alongside the presence of a substantial portion of individuals adopting joint utility reasoning support the idea that perspectives outside of egocentrism may not be secondary (Dale et al., 2018).

It is also possible that individuals who have the capacity to plan jointly may not have had high enough motivation in the task to engage it. Empirical work points to the idea that when interpreting referring expressions, individuals weigh perspectives depending on context (Heller et al., 2016), leading to a division of labor in communication. One factor that could contribute to this division is an estimation of the degree of effort one's partner is exerting (R. D. Hawkins, Gweon, & Goodman, 2021). The debriefing data supports that participants likely believed their partners to be at least somewhat collaborative and the majority of self-reported strategies indicated reasoning about one's partner cooperatively. However, as there was no true interaction between participants, this still may not have been enough motivation to choose jointly efficient actions.

This leaves room for interesting potential future research. Although framed as cooperative, the signaler's responses were pre-programmed. A version of this task that was interactive, bringing subjects in to work together in person or in real-time online, could encourage more robust collaboration. Even more strongly, a version of this experiment where the role of the communicator and listener were not fixed could lead to much stronger preferences for fairness and cooperation. Here, the signaler is in a more powerful position: they have additional information about the target as well as the capacity to force their partner to be responsible for acting by sending a signal. Role reversal in the game could help enforce stronger cooperative norms.

Finally, while these experiments have shown how individuals act when forced to choose between action and belief driven strategies, these strategies are not mutually exclusive. In fact context — and the constraints it provides — likely *accumulates* evidence to resolve ambiguity in linguistic communication (Roy & Mukherjee, 2005). Integration of many simpler contextual heuristics may be a key to fast, flexible, and sparse signaling. Future behavioral research should address how these heuristics interact with each other, which has already been demonstrated to be a theoretically promising approach to communication (Stacy et al., 2021).

CHAPTER 6

Discussion and Directions for Future Work

6.1 How can we say so much with so little?

Even something as effortless as a shared look can be instantly understood as secretive, incredulous, or confused. Humans communicate so much with so little: sparse gestures and simple verbal utterances can often be understood as as complex, sophisticated exchanges. How can this be formalized to build a mechanistic model of human-like communication? The chapters of this dissertation provide a theoretically motivated infrastructure to advance this direction. Here, I take the perspective that intelligence can be reverse-engineered starting from an understanding of the remarkably rich social and physical knowledge exhibited by infants and young toddlers. Long before children master the complexities of language, they are already developing incredibly sophisticated socio-cognitive abilities to support the development of communication. By capturing these underlying mechanisms, I provide the formal structure needed to model overloaded communication at its origin. Specifically, while ToM serves as a widely recognized building block for social intelligence, it is also crucial for the inferential attitude toward communication taken here. Moreover, cooperation plays a critical role in the development of human-unique communication which motivates our Imagined We approach.

Taken together, the chapters of this dissertation integrate and extend existing modeling work to capture overloaded, spontaneous communication. Chapter 2 provides an overview of existing modeling work covering the three major building blocks emphasized in this dissertation: Theory of Mind, cooperation, and communication. Following this, Chapter 3 formally connects two of these ideas: ToM and RSA. This serves to both frame signals in terms of their use in the real world and provide a modeling mechanism to support an inferential perspective on communication. Moreover, it introduces the idea of treating communication as a type of paternalistic helping where the signaler coordinates different perspectives. Next, in Chapter 4, I integrate ToM and RSA with shared agency to introduce the Imaged We model for communication. Previously, in Chapter 3, agents were modeled with unbalanced roles: the signaler acted as an observing helper who could provide information to an actor. However, most settings involve agents who interact in the environment as equal partners. This interaction creates the need for understanding what is mutually known and desired as well as acting on what is mutually efficient. We establish this through a shared agency approach. Finally, in Chapter 5, I provide preliminary behavioral evidence to complement this line of modeling work.

6.2 Broader Insights: Avoiding Traps of Recursion with IW

One major theoretical contribution of adopting an IW framework is that it provides an elegant solution to the problem of recursion involved in many traditional inference models. From the recursion standpoint, selecting as signal involves modeling what I think you think, but also what I think you think I think and so on. Despite the fact that the idea of relying on deep recursion as a means to solve social negotiation has been criticized in cognitive science (Chater, Zeitoun, & Melkonyan, 2022; Sperber & Wilson, 1986; Clark & Marshall, 1981), many modeling approaches do rely on direct partner recursion as a means to coordinate. For example, both BToM and RSA fall into this arena of social recursion. Chater et al. (2022) posit that neither type of model is able to resolve the paradox created from mutual mind-reading. BToM models have been extremely successful at capturing interesting social inferences; however, they sidestep the core challenge of recursion: social interaction. Instead, they model inferences from the perspective of an observer who cannot not intervene in the environment. Thus the target of the observer's inference — the agent interacting in the environment — does not need to mutually reason about or even be aware of this observer. On the other hand, RSA involves direct partner reasoning. However, it does not satisfactorily solve how to coordinate which level of reasoning to adopt or when to stop. The typical method for resolving this is to cut off recursion at a relatively shallow depth which has been justified using behavioral evidence in game-theoretic paradigms which find strategic thinking to be shallow and constrained by working memory (Camerer et al., 2004). However, more recently, depth of recursion has been found to be highly contextual in behavioral economic settings (Georganas, Healy, & Weber, 2015) and much easier and extensive than previously thought when presented implicitly using stories (O'Grady, Kliesch, Smith, & Scott-Phillips, 2015). This indicates the coordination problem may not be trivial. IW offers an alternative to this paradox. While it can be modeled as a recursive process, the most natural stopping point is for all agents to model the same shared We using mutually known, public information. Instead of debating how deep recursion should be, this shifts the focus to the structure of the minds that are being reasoned about. IW reasons about a richer shared mind in a broader context which means it does not need to go as deep.

IW can be thought of as a special type of perspective taking that supersedes any single individual, self or other, which has interesting implications for other processes involved in social cognition. For example, it proposes a cooperative constraint to joint actions such as coordinating how to move a table together, even in the absence of communication. In this sense, instead of anticipating how one's partner might move and how to anticipate and react to that movement for a meshing of plans, the joint goal of moving is simply planned as if agents were already acting together as a single We. Communication acts as a means to course correct when one's partner is not in the ways one's version of We predicted. This is not to say that individuals must always operate under IW, but to argue for it as a mode of reasoning that communicators who have established cooperative motivations can engage in to ease the burden of deep mind-reading.

6.3 Open Questions and Future Directions

While IW offers a novel framework for communication as a means to coordinate perspectives, there are still many open questions and future directions that this modeling approach generates. For example, while I have argued for IW as inspired by a developmental and cognitive perspective, there are cases where communicators still need to engage in more complex reasoning, either to fix cooperation that has gone awry or to establish cooperation and align interests in the first place. One important modeling direction for this framework involves figuring out when and how to recognize you are not "on the same page" as your collaborator and debug this or how to make that initial establishment of a shared We among individuals with imperfectly aligned personal goals and motivations.

Thus far, our formulation of overloaded communication has been highly grounded in the perceptual environment, sensitive to costs under a joint task. In conjunction, I have argued that the core purpose of communication is to coordinate perspectives under uncertainty or asymmetry in knowledge, in this task knowledge of the goal. However, the environment themselves (baring some key piece of knowledge) are fully observable with both agents sharing a top-down perspective. In the future, I could move this into a partially observable domain with different visual perspectives which would lead to beliefs about the environment that could be coordinated in a similar fashion under IW.

Finally, on the behavioral side, many of these theoretical modeling ideas still need rigorous empirical support. While I have provided initial steps toward this in Chapter 5, I focus on preferences between different strategies for resolving overloaded communication. In reality, it is likely that there are many processes and heuristics humans use in conjunction to successfully handle overloading. One next step is to broaden the types of context examined and look at their interaction: how they can build on top of each other to support a fuller inferential account of communication?

APPENDIX A

Pseudocode of IW Algorithm (Chapter 4)

The base case for IW signaler is sending a signal in the signal space under uniform distribution.

Algorithm 1 IW Signaler Algorithm

Inputs:

$goal_t$: true goal;

Ω_s: signal space; Ω_a: action space; Ω_g: goal space;

$JP(g)$: a sequence of actions generated by a joint policy for goal g;

β: rationality parameter;

$C(signal)$: cost of $signal$

$IW_r(signal, \Omega_a, \Omega_g,$JP(.)$, \beta, C(.))$: IW interpretation of a signal, returns IW's goal distribution given a signal

$U($JP(.)$, g)$ = Reward(g) - Cost(JP(g)): utility function as the reward of reaching goal g minus action costs

Output: $decision$: $\{signal \in \Omega_s$, $quit$, $DIY\}$

procedure $IW_s(goal_t, \Omega_s, \Omega_a, \Omega_g, \mathrm{JP}(), \beta, C_s)$

  for $signal \in \Omega_s$ **do**

    for $goal \in \Omega_g$ **do**

      ▷ Signaler's model of how the group should interpret the signal

      $P(goal|signal) = IW_r(signal, \Omega_a, \Omega_g, \mathrm{JP}(), \beta, C_s)[\text{goal}]$

      $P(actions|goal) = \mathbb{1}\{actions = JP(goal)\}$

    end for

    $P(actions|signal) = \sum_{goal \in \Omega_g} P(goal|signal) P(actions|goal)$

    $\mathbb{E}[U(signal, goal_t)] = \mathbb{E}_{P(actions|signal)}[U(actions, goal_t)] - C(signal)$

  end for

end procedure

▷ Compute utility of non-communicative actions:

$U(quit, goal_t) = 0$

$U(DIY, goal_t) = Reward(goal_t) - Cost(\pi_{goal}^{signaler})$ ▷ Utility of signaler reaching goal

$\Omega_d = \{\Omega_s, DIY, quit\}$ ▷ Decision space includes signaling and non-communicative actions

$P(decision|goal_t) = \frac{e^{\beta \mathbb{E}[U(decision, goal_t)]}}{\sum_{d \in \Omega_d} e^{\beta \mathbb{E}[U(d, goal_t)]}}$

Sample a *decision* based on $P(decision|goal_t)$

return *decision*

Algorithm 2 IW Receiver Algorithm

Inputs:

$signal$: signal received;

Ω_s: signal space; Ω_a: action space; Ω_g: goal space;

$JP(g)$: a sequence of actions generated by a joint policy for goal g;

β: rationality parameter;

$IW_s(g, \Omega_s, \Omega_a, \Omega_g, JP(), \beta, C_s)$: Model of how IW signaler selects a signal given goal g

U(JP(.), g) = Reward(g) - Cost(JP(g)): utility function as the reward of reaching goal g minus action costs

Output:

$JP(g_{we})$: The sampled action trajectory to reach the inferred goal g_{we}

1: **procedure** $IW_r(signal, \Omega_a, \Omega_g, JP(), \beta, C_s)$

2: $\triangleright$ Bayes theorem where the likelihood is the IW signaler and $P(goal)$ is assumed uniform across all consistent goals

3: $P(goal|signal) = \frac{(IW_s(goal,\Omega_s,\Omega_a,\Omega_g,JP(),\beta,C_s)[signal])P(goal)}{\sum_{g\in\Omega_g}(IW_s(g,\Omega_s,\Omega_a,\Omega_g,JP(),\beta,C_s)[signal])P(g)}$

4: **end procedure**

5: Sample a g_{we} from $P(goal|signal)$

6: **return** $JP(g_{we})$

References

Adam, C., & Gaudou, B. (2016). BDI agents in social simulations: A survey. *The Knowledge Engineering Review*, *31*(3), 207–238.

Allen, J. F., & Perrault, C. R. (1980). Analyzing intention in utterances. *Artificial Intelligence*, *15*(3), 143–178.

Ashok Kumar, A., Garg, K., & Hawkins, R. (2021). Contextual flexibility guides communication in a cooperative language game. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 43).

Austin, J. L. (1962). *How to do things with words.* Oxford, UK: Oxford University Press.

Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, *1*, 1–10.

Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition*, *113*, 329–349.

Baker, C. L., & Tenenbaum, J. B. (2014). Modeling human plan recognition using Bayesian Theory of Mind. In G. Sukthankar, C. Geib, H. Bui, D. Pynadath, & R. P. Goldman (Eds.), *Plan, activity, and intent recognition: Theory and practice* (pp. 177–204). Elsevier.

Bakker, M., & Wicherts, J. M. (2014). Outlier removal, sum scores, and the inflation of the type I error rate in independent samples t tests: The power of alternatives and recommendations. *Psychological Methods*, *19*(3), 409.

Barr, D. J. (2014). Perspective taking and its impostors in language use: Four patterns of deception. In T. M. Holtgraves (Ed.), *Oxford library of psychology. the oxford handbook of language and social psychology* (pp. 98 – 110)). Oxford, UK: Oxford University Press.

Barth, J., & Call, J. (2006). Tracking and displacement of objects: A series of tasks with great apes (Pan troglodytes, Pan paniscus, Gorilla gorilla, and Pongo pygaeus) and young children (Homo sapiens). *Journal of Experimental Psychology: Animal Behavior Processes*, *32*(3), 239–252.

Berner, C., Brockman, G., Chan, B., Cheung, V., Debiak, P., Dennison, C., ... others (2019). Dota 2 with large scale deep reinforcement learning. *arXiv e-prints*.

Bernstein, D. S., Givan, R., Immerman, N., & Zilberstein, S. (2002). The complexity of decentralized control of Markov decision processes. *Mathematics of Operations Research*, *27*(4), 819–840.

Bohn, M., Tessler, M. H., & Frank, M. C. (2019). Integrating common ground and informativeness in pragmatic word learning. In C. S. Ashok Goel & C. Freksa (Eds.), *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 41, pp. 152–158). Cognitive Science Society.

Bratman, M. (1987). *Intention, plans, and practical reason* (Vol. 10). Cambridge, MA: Harvard University Press.

Bratman, M. (1992). Shared cooperative activity. *The Philosophical Review*, *101*(2), 327–341.

Bratman, M. (2013). *Shared agency: A planning theory of acting together.* Oxford, UK: Oxford University Press.

Bruner, J. (1985). Child's talk: Learning to use language. *Child Language Teaching and Therapy*, *1*(1), 111–114.

Buttelmann, D., Buttelmann, F., Carpenter, M., Call, J., & Tomasello, M. (2017). Great apes distinguish true from false beliefs in an interactive helping task. *PLOS ONE*, *12*(4), e0173793.

Buttelmann, D., Carpenter, M., Call, J., & Tomasello, M. (2007). Enculturated chimpanzees imitate rationally. *Developmental Science*, *10*(4), F31–F38.

Call, J., & Tomasello, M. (2020). *The gestural communication of apes and monkeys.* London, UK: Psychology Press.

Camerer, C. F., Ho, T.-H., & Chong, J.-K. (2004). A cognitive hierarchy model of games. *The Quarterly Journal of Economics*, *119*(3), 861–898.

Carstensen, A., Kon, E., & Regier, T. (2014). Testing a rational account of pragmatic reasoning: The case of spatial language. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 36, p. 2009–2013). Austin, TX: Cognitive Science Society.

Chater, N., Zeitoun, H., & Melkonyan, T. (2022). The paradox of social interaction: Shared intentionality, we-reasoning, and virtual bargaining. *Psychological Review*, *129*(3), 415.

Chomsky, N. (1983). Mental representations. *Syracuse Scholar (1979-1991)*, *4*(2), 1–17.

Clark, H. H. (1992). *Arenas of language use*. Chicago, IL: University of Chicago Press.

Clark, H. H. (1996). *Using language*. Cambridge, MA: Cambridge University Press.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In *Perspectives on socially shared cognition* (pp. 127–149). American Psychological Association.

Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 10 – 63). Cambridge: Cambridge University Press.

Crandall, J. W. (2020). When autonomous agents model other agents: An appeal for altered judgment coupled with mouths, ears, and a little more tape. *Artificial Intelligence*, *280*.

Dale, R., Galati, A., Alviar, C., Contreras Kallens, P., Ramirez-Aristizabal, A. G., Tabatabaeian, M., & Vinson, D. W. (2018). Interacting timescales in perspective-taking. *Frontiers in Psychology*, *9*, 1278.

Degen, J., Hawkins, R., Graf, C., Kreiss, E., & Goodman, N. (2020). When redundancy is useful: A Bayesian approach to "overinformative" referring expressions. *Psychological Review*, *127*(4), 591.

Dennett, D. C. (1987). *The intentional stance*. Cambridge, MA: MIT Press.

De Ruiter, J. P., Noordzij, M., Newman-Norlund, S., Hagoort, P., & Toni, I. (2007). On the origin of intentions. *Attention & Performance XXII*, 593–610.

Doshi, P., Gmytrasiewicz, P., & Durfee, E. (2020). Recursively modeling other agents for decision making: A research perspective. *Artificial Intelligence*, *279*, 103202.

Duguid, S., Wyman, E., Bullinger, A. F., Herfurth-Majstorovic, K., & Tomasello, M. (2014). Coordination strategies of chimpanzees and human children in a Stag Hunt game. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1796).

Emery, N. J., & Clayton, N. S. (2004). The mentality of crows: Convergent evolution of intelligence in corvids and apes. *Science*, *306*(5703), 1903–1907.

Engelmann, J. M., & Tomasello, M. (2019). Children's sense of fairness as equal respect. *Trends in Cognitive Sciences*, *23*(6), 454–463.

Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004). Perspective taking as egocentric anchoring and adjustment. *Journal of Personality and Social Psychology*, *87*(3), 327.

Fan, X., & Yen, J. (2004). Modeling and simulating human teamwork behaviors using intelligent agents. *Physics of Life Reviews*, *1*(3), 173–201.

Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, *336*(6084), 998.

Fussell, S. R., & Krauss, R. M. (1992). Coordination of knowledge in communication: Effects of speakers' assumptions about what others know. *Journal of Personality and Social Psychology*, *62*(3), 378 – 391.

Gallotti, M., & Frith, C. D. (2013). Social cognition in the we-mode. *Trends in Cognitive Sciences*, *17*(4), 160–165.

Ganea, P. A., & Saylor, M. M. (2007). Infants' use of shared linguistic information to clarify ambiguous requests. *Child Development*, *78*(2), 493–502.

Gao, T., Baker, C. L., Tang, N., Xu, H., & Tenenbaum, J. B. (2019). The cognitive architecture of perceived animacy: Intention, attention, and memory. *Cognitive Science*, *43*(8), e12775.

Georganas, S., Healy, P. J., & Weber, R. A. (2015). On the persistence of strategic sophistication. *Journal of Economic Theory*, *159*, 369–400.

Georgeff, M., Pell, B., Pollack, M., Tambe, M., & Wooldridge, M. (1998). The belief-desire-intention model of agency. In *International workshop on agent theories, architectures, and languages* (pp. 1–10).

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*(2), 165–193.

Gilbert, M. (1992). *On social facts.* Princeton, NJ: Princeton University Press.

Gilbert, M. (1999). Obligation and joint commitment. *Utilitas: A Journal of Utilitarian Studies*, *11*(2), 143–163.

Gilbert, M. (2013). *Joint commitment: How we make the social world.* Oxford, UK: Oxford University Press.

Gmytrasiewicz, P. (2020). How to do things with words: A Bayesian approach. *Journal of Artificial Intelligence Research*, *68*, 753–776.

Gmytrasiewicz, P. J., & Doshi, P. (2005). A framework for sequential planning in multi-agent settings. *Journal of Artificial Intelligence Research*, *24*, 49–79.

Gmytrasiewicz, P. J., & Durfee, E. H. (2001). Rational communication in multi-agent environments. *Autonomous Agents and Multi-Agent Systems*, *4*(3), 233–272.

Goodman, N. D., & Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, *20*, 818–829.

Goodman, N. D., & Stuhlmüller, A. (2013). Knowledge and implicature: Modeling language understanding as social cognition. *Topics in Cognitive Science*, *5*(1), 173–184.

Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories.* Cambridge, MA: MIT Press.

Gräfenhain, M., Behne, T., Carpenter, M., & Tomasello, M. (2009). Young children's understanding of joint commitments. *Developmental Psychology*, *45*(5), 1430–1443.

Grice, H. P. (1975). Logic and conversation. In P. Cole & M. J (Eds.), *Syntax and sematics*

(Vol. 3, p. 41 - 58). New York: Academic Press.

Grosse, G., Moll, H., & Tomasello, M. (2010). 21-month-olds understand the cooperative logic of requests. *Journal of Pragmatics*, *42*(12), 3377–3383.

Grosz, B. J., & Hunsberger, L. (2006). The dynamics of intention in collaborative activity. *Cognitive Systems Research*, *7*(2-3), 259–272.

Grosz, B. J., & Kraus, S. (1996). Collaborative plans for complex group action. *Artificial Intelligence*, *86*(2), 269–357.

Hamann, K., Warneken, F., Greenberg, J. R., & Tomasello, M. (2011). Collaboration encourages equal sharing in children but not in chimpanzees. *Nature*, *476*(7360), 328–331.

Hamann, K., Warneken, F., & Tomasello, M. (2012). Children's developing commitments to joint goals. *Child Development*, *83*(1), 137–145.

Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language*, *49*(1), 43–61.

Hare, B., Call, J., Agnetta, B., & Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour*, *59*(4), 771–785.

Hare, B., Call, J., & Tomasello, M. (2006). Chimpanzees deceive a human competitor by hiding. *Cognition*, *101*(3), 495–514.

Hare, B., & Tomasello, M. (2004). Chimpanzees are more skilful in competitive than in cooperative cognitive tasks. *Animal Behaviour*, *68*(3), 571–581.

Havrylov, S., & Titov, I. (2017). Emergence of language with multi-agent games: Learning to communicate with sequences of symbols. *Advances in Neural Information Processing Systems*, *30*.

Hawkins, R. D., Gweon, H., & Goodman, N. D. (2021). The division of labor in communication: Speakers help listeners account for asymmetries in visual perspective. *Cognitive Science*, *45*(3).

Hawkins, R. X., Frank, M., & Goodman, N. D. (2017). Convention-formation in iterated reference games. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 39).

Hayashi, A., Ruiken, D., Hasegawa, T., & Goerick, C. (2020). Reasoning about uncertain parameters and agent behaviors through encoded experiences and belief planning. *Artificial Intelligence*, *280*.

Heller, D., Parisien, C., & Stevenson, S. (2016). Perspective-taking behavior as the probabilistic weighing of multiple domains. *Cognition*, *149*, 104–120.

Heritage, J. (1984). *Garfinkel and ethnomethodology.* Cambridge, UK: Polity Press.

Ho, M. K., Cushman, F., Littman, M. L., & Austerweil, J. L. (2021). Communication in action: Planning and interpreting communicative demonstrations. *Journal of Experimental Psychology: General*.

Ho, M. K., Littman, M., MacGlashan, J., Cushman, F., & Austerweil, J. L. (2016). Showing versus doing: Teaching by demonstration. *Advances in Neural Information Processing Systems*, *29*.

Hunt, G. R. (1996). Manufacture and use of hook-tools by New Caledonian crows. *Nature*, *379*(6562), 249–251.

Inoue, S., & Matsuzawa, T. (2007). Working memory of numerals in chimpanzees. *Current Biology*, *17*(23), R1004–R1005.

Isaacs, E. A., & Clark, H. H. (1987). References in conversation between experts and novices. *Journal of Experimental Psychology: General*, *116*(1), 26 – 37.

Jacobsson, F., Johannesson, M., & Borgquist, L. (2007). Is altruism paternalistic? *The Economic Journal*, *117*(520), 761–781.

Jara-Ettinger, J., Floyd, S., Huey, H., Tenenbaum, J. B., & Schulz, L. E. (2020). Social pragmatics: Preschoolers rely on commonsense psychology to resolve referential underspecification. *Child Development*, *91*(4), 1135–1149.

Jara-Ettinger, J., Gweon, H., Schulz, L. E., & Tenenbaum, J. B. (2016). The naïve utility calculus: Computational principles underlying commonsense psychology. *Trends in Cognitive Sciences*, *20*(8), 589–604.

Jennings, N. R. (1995). Controlling cooperative problem solving in industrial multi-agent systems using joint intentions. *Artificial Intelligence*, *75*(2), 195–240.

Jiang, K., Stacy, S., Dahmani, A. L., Jiang, B., Rossano, F., Zhu, Y., & Gao, T. (2022). What is the point? a Theory of Mind model of relevance. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 44).

Jiang, K., Stacy, S., Wei, C., Chan, A., Rossano, F., Zhu, Y., & Gao, T. (2021). Individual vs. joint perception: A pragmatic model of pointing as communicative Smithian helping. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 43).

Kaelbling, L. P., Littman, M. L., & Cassandra, A. R. (1998). Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, *101*(1-2), 99–134.

Kao, J., Bergen, L., & Goodman, N. (2014). Formalizing the pragmatics of metaphor understanding. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 36).

Keysar, B., Barr, D. J., & Horton, W. S. (1998). The egocentric basis of language use: Insights from a processing approach. *Current directions in psychological science*, *7*(2), 46–49.

Kleiman-Weiner, M., Ho, M. K., Austerweil, J. L., Littman, M. L., & Tenenbaum, J. B. (2016). Coordinate to cooperate or compete: Abstract goals and joint intentions in social interaction. In D. M. . J. T. A. Papafragou D. Grodner (Ed.), *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 38, pp. 1679 –1684). Cognitive Science Society.

Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2016). Great apes anticipate that other individuals will act according to false beliefs. *Science*, *354*(6308), 110–114.

Lagoudakis, M. G., & Parr, R. (2002). Learning in zero-sum team Markov games using factored value functions. *Advances in Neural Information Processing Systems*, *15*.

Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2017). Building machines that learn and think like people. *Behavioral and Brain Sciences*, *40*.

Lassiter, D., & Goodman, N. D. (2017). Adjectival vagueness in a Bayesian model of interpretation. *Synthese*, *194*(10), 3801–3836.

Levelt, W. (1989). *Speaking: From intention to articulation.* Cambridge, MA: MIT Press.

Levelt, W. J. (1983). Monitoring and self-repair in speech. *Cognition*, *14*(1), 41–104.

Levesque, H. J., Cohen, P. R., & Nunes, J. H. (1990). On acting together. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 8, pp. 94–99).

Levinson, S. (1983). *Pragmatics.* Cambridge, UK: Cambridge University Press.

Lewis, D. (1969). *Convention: A philosophical study.* New York, NY: John Wiley & Sons.

Liebal, K., Behne, T., Carpenter, M., & Tomasello, M. (2009). Infants use shared experience to interpret pointing gestures. *Developmental Science*, *12*(2), 264–271.

Liszkowski, U., Carpenter, M., & Tomasello, M. (2007). Pointing out new news, old news, and absent referents at 12 months of age. *Developmental Science*, *10*(2), F1–F7.

Liszkowski, U., Schäfer, M., Carpenter, M., & Tomasello, M. (2009). Prelinguistic infants, but not chimpanzees, communicate about absent entities. *Psychological Science*, *20*(5), 654–660.

Liu, S., Ullman, T. D., Tenenbaum, J. B., & Spelke, E. S. (2017). Ten-month-old infants infer the value of goals from the costs of actions. *Science*, *358*(6366), 1038–1041.

Luce, R. D. (1959). On the possible psychophysical laws. *Psychological Review*, *66*, 81–95.

Martin, A., Lin, K., & Olson, K. R. (2016). What you want versus what's good for you: Paternalistic motivation in children's helping behavior. *Child Development*, *87*(6), 1739–1746.

McCarthy, W. P., Hawkins, R., Wang, H., Holdaway, C., & Fan, J. E. (2021). Learning to communicate about shared procedural abstractions. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 43).

Melis, A. P., Call, J., & Tomasello, M. (2006). Chimpanzees (Pan troglodytes) conceal visual and auditory information from others. *Journal of Comparative Psychology*, *120*(2), 154–162.

Melis, A. P., Schneider, A.-C., & Tomasello, M. (2011). Chimpanzees, Pan troglodytes, share food in the same way after collaborative and individual food acquisition. *Animal Behaviour*, *82*(3), 485–493.

Misyak, J., Melkonyan, T., Zeitoun, H., & Chater, N. (2014). Unwritten rules: virtual bargaining underpins social interaction, culture, and society. *Trends in Cognitive Sciences*, *18*(10), 512–519.

Misyak, J., Noguchi, T., & Chater, N. (2016). Instantaneous conventions: The emergence of flexible communicative signals. *Psychological Science*, *27*(12), 1550–1561.

Moll, H., Richter, N., Carpenter, M., & Tomasello, M. (2008). Fourteen-month-olds know what "we" have shared in a special way. *Infancy*, *13*(1), 90–101.

Nagel, T. (1989). *The view from nowhere.* Oxford, UK: Oxford University Press.

Newell, A., Shaw, J. C., & Simon, H. A. (1957). Empirical explorations of the logic theory machine: A case study in heuristic. In *Papers presented at the February 26-28, 1957, Western Joint Computer Conference: Techniques for Reliability* (pp. 218–230).

Norling, E. (2004). Folk psychology for human modelling: Extending the BDI paradigm. In *Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems* (Vol. 1, pp. 202–209).

Oliehoek, F. A., & Amato, C. (2016). *A concise introduction to decentralized POMDPs.* Springer International Publishing.

O'Grady, C., Kliesch, C., Smith, K., & Scott-Phillips, T. C. (2015). The ease and extent of recursive mindreading, across implicit and explicit tasks. *Evolution and Human Behavior*, *36*(4), 313–322.

Pantelis, P. C., Baker, C. L., Cholewiak, S. A., Sanik, K., Weinstein, A., Wu, C.-C., ... Feldman, J. (2014). Inferring the intentional states of autonomous virtual agents. *Cognition*, *130*(3), 360–379.

Pinker, S. (2007 [1994]). *The language instinct: How the mind creates language.* New York, NY: Haper Perennial Modern Classics.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a Theory of Mind? *Behavioral and Brain Sciences*, *1*(4), 515–526.

Qing, C., & Franke, M. (2015). Variations on a Bayesian theme: Comparing Bayesian models of referential reasoning. In H. Zeevat & H.-C. Schmitz (Eds.), *Bayesian natural language semantics and pragmatics* (pp. 201–220). Springer.

Rao, A. S., & Georgeff, M. P. (1995). BDI agents: From theory to practice. In *Proceedings of the International Conference on Multiagent Systems* (Vol. 95, pp. 312–319).

Regan, P. M., Frank, R. W., & Aydin, A. (2009). Diplomatic interventions and civil war: A new dataset. *Journal of Peace Research*, *46*(1), 135–146.

Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). The art of conversation is coordination. *Psychological Science*, *18*(5), 407–413.

Roy, D., & Mukherjee, N. (2005). Towards situated speech understanding: Visual context priming of language models. *Computer Speech & Language*, *19*(2), 227–248.

Russell, S. (2019). *Human compatible: Artificial intelligence and the problem of control.* London, UK: Penguin.

Sacks, H. (1985). The inference-making machine: Notes on observability. *Handbook of Discourse Analysis*, *3*, 13–23.

Sayin, M., Zhang, K., Leslie, D., Basar, T., & Ozdaglar, A. (2021). Decentralized Q-learning in zero-sum Markov games. *Advances in Neural Information Processing Systems*, *34*.

Schelling, T. C. (1960). *The strategy of conflict.* Cambridge, MA: Harvard University Press.

Schiffer, S. (1972). Meaning.

Schweikard, D. P., & Schmid, H. B. (2021). Collective intentionality. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy.* Metaphysics Research Lab, Stanford University.

Searle, J. R. (1990). Collective intentions and actions. In *Intentions in communication.*

Selten, R., & Warglien, M. (2007). The emergence of simple languages in an experimental coordination game. *Proceedings of the National Academy of Sciences*, *104*(18), 7361–7366.

Shalev-Shwartz, S., Shammah, S., & Shashua, A. (2016). Safe, multi-agent, reinforcement learning for autonomous driving. *arXiv preprint arXiv:1610.03295*.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(3), 379–423.

Shum, M., Kleiman-Weiner, M., Littman, M. L., & Tenenbaum, J. B. (2019). Theory of Minds: Understanding behavior in groups through inverse planning. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 33, pp. 6163–6170).

Sibicky, M. E., Schroeder, D. A., & Dovidio, J. F. (1995). Empathy and helping: Considering the consequences of intervention. *Basic and Applied Social Psychology*, *16*(4), 435–453.

Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., … others (2017). Mastering the game of Go without human knowledge. *Nature*, *550*(7676), 354–359.

Simon, H. A., & Newell, A. (1962). Computer simulation of human thinking and problem solving. *Monographs of the Society for Research in Child Development*, *27*(2), 137–150.

Singh, D., Padgham, L., & Logan, B. (2016). Integrating BDI agents with agent-based simulation platforms. *Autonomous Agents and Multi-Agent Systems*, *30*(6), 1050–1071.

Siposova, B., Tomasello, M., & Carpenter, M. (2018). Communicative eye contact signals a commitment to cooperate for young children. *Cognition*, *179*, 192–201.

Skyrms, B. (2010). *Signals: Evolution, learning, and information.* Oxford, UK: Oxford University Press.

Smith, A. (2010 [1759]). *The theory of moral sentiments.* London, UK: Penguin.

Spaan, M. T., Gordon, G. J., & Vlassis, N. (2006). Decentralized planning under uncertainty for teams of communicating agents. In *Proceedings of the International Joint*

Conference on Autonomous Agents and Multiagent Systems (pp. 249–256).

Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science*, *10*(1), 89–96.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Cambridge, MA: Harvard University Press.

Stacy, S., Li, C., Zhao, M., Yun, Y., Zhao, Q., Kleiman-Weiner, M., & Gao, T. (2021). Modeling communication to coordinate perspectives in cooperation. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 43).

Stacy, S., Parab, A., Kleiman-Weiner, M., & Gao, T. (2022). Overloaded communication as paternalistic helping. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 44).

Stacy, S., Yun, Y., Potter, M., Moskowitz, N., & Gao, T. (2022). No such thing as the average listener: Belief-driven versus action-driven strategies in signaling. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 44).

Sumers, T., Hawkins, R., Ho, M. K., & Griffiths, T. (2021). Extending rational models of communication from beliefs to actions. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 43).

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. Cambridge, MA: MIT press.

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2007). *Using multivariate statistics* (Vol. 5). Boston, MA: Pearson.

Tambe, M. (1997). Towards flexible teamwork. *Journal of Artificial Intelligence Research*, *7*, 83–124.

Tang, N., Gong, S., Zhao, M., Gu, C., Zhou, J., Shen, M., & Gao, T. (2022). Exploring an imagined "we" in human collective hunting: Joint commitment within shared intentionality. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 44).

Tang, N., Stacy, S., Zhao, M., Marquez, G., & Gao, T. (2020). Bootstrapping an imagined we for cooperation. In Y. X. . B. A. S. Denison. M. Mack (Ed.), *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 42, pp. 2453–2458). Cognitive Science Society.

Tomasello, M. (2000). The social-pragmatic theory of word learning. *Pragmatics*, *10*(4), 401–414.

Tomasello, M. (2010). *Origins of human communication.* Cambridge, MA: MIT Press.

Tomasello, M. (2019). *Becoming human.* Cambridge, MA: Harvard University Press.

Tomasello, M., Call, J., & Gluckman, A. (1997). Comprehension of novel communicative signs by apes and human children. *Child Development*, 1067–1080.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, *28*(5), 675–691.

Tomasello, M., Carpenter, M., & Hobson, R. P. (2005). The emergence of social cognition in three young chimpanzees. *Monographs of the Society for Research in Child Development*, i–152.

Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development*, *78*(3), 705–722.

Tomasello, M., Davis-Dasilva, M., CamaK, L., & Bard, K. (1987). Observational learning of tool-use by young chimpanzees. *Human Evolution*, *2*(2), 175–183.

Tomasello, M., & Haberl, K. (2003). Understanding attention: 12-and 18-month-olds know what is new for other persons. *Developmental Psychology*, *39*(5), 906–912.

Török, G., Pomiechowska, B., Csibra, G., & Sebanz, N. (2019). Rationality in joint action: Maximizing coefficiency in coordination. *Psychological Science*, *30*(6), 930–941.

Townsend, J. T. (1992). On the proper scales for reaction time. *Cognition, information processing, and psychophysics: Basic issues*, 105–120.

Tuomela, R. (2007). *The philosophy of sociality: The shared point of view.* Oxford, UK:

Oxford University Press.

Turing, A. M. (2009 [1950]). Computing machinery and intelligence. In R. Epstein, G. Roberts, & G. Beber (Eds.), *Parsing the turing test* (pp. 23–65). Springer.

Ullman, T., Baker, C., Macindoe, O., Evans, O., Goodman, N., & Tenenbaum, J. (2009). Help or hinder: Bayesian models of social goal inference. In *Advances in Neural Information Processing Systems* (Vol. 22, pp. 1874–1882).

Ullman, T., Spelke, E., Battaglia, P., & Tenenbaum, J. B. (2017). Mind games: Game engines as an architecture for intuitive physics. *Trends in Cognitive Sciences*, *21*, 649–665.

Vaish, A., Carpenter, M., & Tomasello, M. (2016). The early emergence of guilt-motivated prosocial behavior. *Child Development*, *87*(6), 1772–1782.

Valiant, L. G. (1984). A theory of the learnable. *Communications of the ACM*, *27*(11), 1134–1142.

Van Rooy, R. (2003). Questioning to resolve decision problems. *Linguistics and Philosophy*, *26*(6), 727–763.

Vinyals, O., Babuschkin, I., Czarnecki, W. M., Mathieu, M., Dudzik, A., Chung, J., ... others (2019). Grandmaster level in starcraft II using multi-agent reinforcement learning. *Nature*, *575*(7782), 350–354.

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes.

Warneken, F., Chen, F., & Tomasello, M. (2006). Cooperative activities in young children and chimpanzees. *Child Development*, *77*(3), 640–663.

Warneken, F., Hare, B., Melis, A. P., Hanus, D., & Tomasello, M. (2007). Spontaneous altruism by chimpanzees and young children. *PLOS Biology*, *5*(7), e184.

Warneken, F., Lohse, K., Melis, A. P., & Tomasello, M. (2011). Young children share the spoils after collaboration. *Psychological Science*, *22*(2), 267–273.

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science*, *311*(5765), 1301–1303.

Wellman, H. M. (1992). *The child's Theory of Mind*. Cambridge, MA: The MIT Press.

Wharton, T. (2003). Natural pragmatics and natural codes. *Mind & Language*, *18*(5), 447–477.

Whorf, B. L. (1956). *Language, thought, and reality: Selected writings of Benjamin Lee Whorf* (J. Carroll, Ed.). Cambridge, MA: MIT Press.

Wittgenstein, L. (1953). *Philosophical investigations*. Basil Blackwell.

Wobber, V., Herrmann, E., Hare, B., Wrangham, R., & Tomasello, M. (2014). Differences in the early cognitive development of children and great apes. *Developmental Psychobiology*, *56*(3), 547–573.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*(1), 1–34.

Wu, S. A., Wang, R. E., Evans, J. A., Tenenbaum, J. B., Parkes, D. C., & Kleiman-Weiner, M. (2021). Too many cooks: Bayesian inference for coordinating multi-agent collaboration. *Topics in Cognitive Science*, *13*(2), 414–432.

Wyman, E., Rakoczy, H., & Tomasello, M. (2013). Non-verbal communication enables children's coordination in a "stag hunt" game. *European Journal of Developmental Psychology*, *10*(5), 597–610.

Xie, Q., Chen, Y., Wang, Z., & Yang, Z. (2020). Learning zero-sum simultaneous-move Markov games using function approximation and correlated equilibrium. In *Conference on Learning Theory* (pp. 3674–3682).

Xuan, P., Lesser, V., & Zilberstein, S. (2001). Communication decisions in multi-agent cooperation: Model and experiments. In *Proceedings of the fifth international conference on autonomous agents* (pp. 616–623).

Yamamoto, S., Humle, T., & Tanaka, M. (2012). Chimpanzees' flexible targeted helping based on an understanding of conspecifics' goals. *Proceedings of the National Academy of Sciences*, *109*(9), 3588–3592.

Yob, G. (1975). Hunt the wumpus. *Creative Computing*, 51–54.

Yoon, E. J., Tessler, M. H., Goodman, N. D., & Frank, M. C. (2017). "I won't lie, it wasn't amazing": Modeling polite indirect speech. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 39).

Yuan, A. X., Monroe, W., Bai, Y., & Kushman, N. (2018). Understanding the Rational Speech Act model. In J. Z. Chuck Kalish Martina Rau & T. Rogers (Eds.), *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 40, pp. 2759 – 2764). Cognitive Science Society.

Zhang, K., Kakade, S., Basar, T., & Yang, L. (2020). Model-based multi-agent RL in zero-sum Markov games with near-optimal sample complexity. *Advances in Neural Information Processing Systems*, *33*, 1166–1178.

Zhang, K., Yang, Z., Liu, H., Zhang, T., & Basar, T. (2018). Fully decentralized multi-agent reinforcement learning with networked agents. In *International Conference on Machine Learning* (pp. 5872–5881).

Zhao, M., Tang, N., Dahmani, A. L., Zhu, Y., Rossano, F., & Gao, T. (2022). Sharing rewards undermines coordinated hunting. *Journal of Computational Biology*.

Zhu, Y., Gao, T., Fan, L., Huang, S., Edmonds, M., Liu, H., ... Zhu, S.-C. (2020). Dark, beyond deep: A paradigm shift to cognitive AI with humanlike common sense. *Engineering*, *6*(3), 310–345.