A Rational Choice Framework for Collective Behavior

by

Peter M. Krafft

A.A., Massasoit Community College (2007)
B.S., University of Massachusetts Amherst (2010)
M.S., University of Massachusetts Amherst (2012)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2017

Author .....
Department of Electrical Engineering and Computer Science
August 18, 2017

Certified by .....
Joshua B. Tenenbaum
Professor of Brain and Cognitive Sciences
Thesis Supervisor

Certified by .....
Alex "Sandy" Pentland
Toshiba Professor of Media Arts and Sciences
Thesis Supervisor

Accepted by .....
Leslie A. Kolodziejski
Professor of Electrical Engineering and Computer Science
Chair, Department Committee on Graduate Students

A Rational Choice Framework for Collective Behavior

by

Peter M. Krafft

Submitted to the Department of Electrical Engineering and Computer Science
on August 18, 2017, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

As the world becomes increasingly digitally mediated, people can more and more easily form groups, teams, and communities around shared interests and goals. Yet there is a constant struggle across forms of social organization to maintain stability and coherency in the face of disparate individual experiences and agendas. When are collectives able to function and thrive despite these challenges? In this thesis I propose a theoretical framework for reasoning about collective intelligence—the ability of people to accomplish their shared goals together. A simple result from the literature on multiagent systems suggests that strong general collective intelligence in the form of "rational group agency" arises from three conditions: aligned utilities, accurate shared beliefs, and coordinated actions. However, achieving these conditions can be difficult, as evidenced by impossibility results related to each condition from the literature on social choice, belief aggregation, and distributed systems. The theoretical framework I propose serves as a point of inspiration to study how human groups address these difficulties. To this end, I develop computational models of facets of human collective intelligence, and test these models in specific case studies. The models I introduce suggest distributed Bayesian inference as a framework for understanding shared belief formation, and also show that people can overcome other difficult computational challenges associated with achieving rational group agency, including balancing the group "exploration versus exploitation dilemma" for information gathering and inferring levels of "common p-belief" to coordinate actions.

Thesis Supervisor: Joshua B. Tenenbaum
Title: Professor of Brain and Cognitive Sciences

Thesis Supervisor: Alex "Sandy" Pentland
Title: Toshiba Professor of Media Arts and Sciences

Acknowledgments

As I think is more than appropriate, this acknowledgments section might be longer than the actual thesis itself.

I will start by thanking my advisors, Sandy and Josh. Sandy has been unimaginably patient and supportive. There have been times where I have disappeared for the better part of several months without seeing Sandy, and Sandy has always trusted me without complaint during these times. I believe my work is all the better for it. Sandy also allowed me to have a huge amount of freedom in what work that I chose to pursue, and he was always encouraging about the projects I brought to him. Sandy applied just enough pressure for me to know that he was there and that I should be progressing, while not applying so much pressure as to cause undue stress. The scientific problems that I ultimately ended up working on were also heavily influenced by conversations with Sandy, by reading his past work, and by belonging to the community he helped to build. Sandy will continue to be a role model to me as someone who maintains an impressive level of awareness of a startling number of technological advances and who engages with a huge range of different types of people, from academia, industry, government, and beyond.

Josh has influenced me profoundly in different ways. Whereas Sandy was by and large uniformly encouraging, Josh provided a more critical eye. Josh has an unwavering commitment to a high degree of scientific rigor and an unapologetic commitment to his fine taste in research questions. The basic questions that I have chosen to pursue and the most interesting parts of the work that I have done were heavily influenced by Josh's ways of thinking. I have an immense respect for Josh, and I remain in awe at his boundless energy and the facility with which he engages in a range of intellectual discussion. I also feel a personal connection to Josh that I think we both recognize but that we have not yet had enough time or occasion to fully realize. I hope we continue to collaborate and that some day we can more deeply explore our personal perspectives and shared interests.

I would next like to thank the two other members of my committee, Leslie Pack

Kaelbling and John Tsitsiklis. I was admitted to MIT as an AI person, and because of that I have had the pleasure of interacting with Leslie many times over the course of my PhD. Leslie has been invariably welcoming and generous, always willing to talk when I have needed direction in my research or professional advice and always providing useful feedback in both avenues. John I met later in my program, and I feel fortunate to have finally had that opportunity. When I approached John about joining my committee, I told him about my work, and he had many insightful remarks. Since then John has continued to offer constructive criticism while also remaining open-minded. Thank you both for joining my committee on somewhat short notice and being so exceptionally accommodating.

I have also been fortunate to have had another mentor during my PhD in Barbara Grosz. I initially reached out to Barbara to talk about some research ideas. Ultimately we ended up organizing two reading groups and one class together. Discussing readings with Barbara was always stimulating, and I learned a great deal about teaching as her teaching assistant. Barbara had seemingly endless patience with the students in our class, yet simultaneously demanded that the students meet a high standard of excellence for their coursework and conduct. Outside of the classroom, I cherish Barbara as a friend and always look forward to our conversations.

I must also thank Hanna Wallach. Hanna might have played the single greatest role in shaping my graduate career. Academically, Hanna shepherded my transition to Bayesian machine learning and introduced me to computational social science. Professionally, Hanna was a tireless advocate, and her support was central to my success in PhD applications. Personally, Hanna stands as a role model for how to simultaneously be an accomplished academic and an awesome person with many interesting friends, hobbies, and preoccupations. I look forward to hanging out in New York. :)

George Konidaris also played a huge role in my graduate career. I worked with George and Andy Barto as an undergraduate at UMass. I probably would not have gone to graduate school if not for George's encouragement, and I think I can honestly say that I am happy with my decision to go. George also alerted me to Hanna's

work while I was at UMass, and to Josh's work when I was thinking about where to apply for grad school. I still enjoy catching up with George at conferences whenever I see him, and I am honored to have had him as a mentor. And I'm finally over my bitterness that he didn't like my idea of making a crowdsourced reinforcement learning artist. Turns out applying deep learning to Go was also a good idea. Sorry I kind of dropped the ball on that one.

Andy Barto was an important mentor to me in his own right as well. I always felt that Andy was in my corner. At the first conference I attended, I hadn't known that I was supposed to register, and Andy gave me his badge. I felt so cool the whole time I wore it. As I progressed to grad school applications, Andy had my back again. I still see Andy now and then, and I feel a rare connection to him.

Farshid Hajir was the first person who I tried to do research with, and was also just an all-around awesome and hip dude. I initially met Farshid super randomly while I was walking in between classes at UMass. He had apparently been planning to meet me, though, because immediately at that chance encounter he invited me to his summer research program. At that program I met some of my best friends from college, and overall had a blast. Research has literally never been so fun.

I spent a lot of time with Michael Lavine as an undergraduate. Prof. Lavine holds the distinction of being the one to have brainwashed me into a Bayesian statistician. In addition to the research that I did with him, I still recall the conversations we had about life. I think Prof. Lavine probably saw me at my most angsty in college, and he helped me keep it real. At a time when I felt like grad school was the only thing I could do with the next few years of my life, Prof. Lavine forced me to make the decision to do grad school carefully.

I have had many other mentors over the years as well, all of whom are special to me in one way or another for the immense impact they have had on my life at the time I knew them: my fourth grade teacher Helen Ann Civian for making me feel special, Prof. Janey at Massasoit for encouraging all his students to go as far in math as they can, Mr. Sarro for giving me a B in history, Dick Bogartz for showing me that stats is not boring, Andrew Cohen for having me in his psychology lab, Ann Schwartz

for happily dealing with my shit in her linear algebra class, and Sridhar Mahadevan for showing me another side of machine learning.

Next, to my family. My mother, who always encouraged me to do what makes me happy. My father, from whom I have learned so much about life. Luke, who has profoundly shaped my interests, politics, and world view. Max, who sparked a love of music and set the bar for academic achievement. Johanna, who is endlessly caring. Bob, for his kindness and good humor.

And, to my friends and colleagues. I am blessed to have too many of you to mention by name, but I appreciate you all. I'll give a couple very special randomly chosen shout-outs, though. Max Siegel has been a great friend, a moral compass, and a vociferous, insatiable critic. Thanks, Max. Thanks to Been, for having a huge and lasting positive impact on my quality of life, from my sleep schedule to my work-life balance. Nikete, you inspire me with your outlook on life and your lifestyle. Robert X.D., you're an amazing collaborator. Elisa, what a fortunate occasion it was to meet you at that CrowdCamp. Ankur, I look forward to finishing our paper together. Abdullah, you're a prince among men. Dhaval, it's been lovely getting to know you better. Kelsey, I wish we'd gotten to hang out more. Box Wine, the Druid crew, Karen Huang, all my Cambridge party people: you've made the last year of my PhD a great year. Young Ji, because you complained about not getting a shout-out :p. Tommy and Julia, it was a blast grinding on thesis-writing together. Thor, Pools, Jooner: looking forward to being reunited.

Finally, I regret to say that I am sure there is someone very important to me who I have forgotten to thank. I can only offer an apology for my forgetfulness. Know that I still appreciate you. I owe you dinner.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

A response I often get when I say that I study collective intelligence is something along the lines of "what a peculiar thing to study; there isn't a lot of that going around". Indeed, there are countless examples of collective failure—from market bubbles and crashes to the spread of false rumors. At the same time, society pushes forward, with technological, scientific, and cultural changes leading to ever-increasing complexity in our world; and many communities, organizations, and institutions persist and better themselves over extended periods of time—from the scientific community to the denizens of Wikipedia. While the failures are notable, collective intelligence abounds.

Granted, in a way, my interlocutors are still right to be surprised. Human society, by and large, is a gigantic mess. People haphazardly stumble through the world with inevitably partial understandings of the state of things. We can only react to our surroundings and do the best we can. Members of non-familial groups are only loosely genetically related and therefore have little biological reason to be linked in their actions. People have differing, sometimes incommensurable, goals and beliefs. We each display remarkably complex and heterogeneous behavioral patterns that have no a priori reason to mesh well together. We lack a centralized controller of individual action to help us coordinate. Even laws and social norms only create soft constraints on individual agency. There is nothing that holds groups together except our will, and the structures and mechanisms we create. The ability of people to spontaneously

form coherent groups amid this chaos is remarkable.

The central questions of this thesis are: How can we formally characterize the fundamental challenges that goal-directed groups face in achieving effective collective action, and how do people overcome these challenges? To answer these questions, I take a computational perspective. I argue that rational group agency arises from three processes—alignment of preferences, beliefs, and actions—each of which is associated with particular challenges. I present computational models that reveal how people achieve distributed information processing and consensus mechanisms to address these challenges. The central thesis these claims support is that many aspects of human collective behavior function to solve the fundamental problems associated with the establishment of collective agency.

1.1 The Philosophy of the Approach

To speak of coherent group action is to suppose that there is regular structure in the behavior of certain groups that can be understood in its own right, perhaps even mathematically modeled, at the group level. To suppose these macro-level regularities exist does not diminish the importance of the individual behaviors that necessarily constitute group behavior.

At the fore of this thesis is this tension in the complex relations and interactions between individual behavior and group behavior. Wrestling with multiple levels of analysis is inevitable as a student of social systems. Individuals determine the behavior of groups, but the aggregate level can also at times be understood in isolation, and is sometimes understood more clearly in isolation. A notable example is simultaneous scientific discovery. The discoveries that are made in a particular period of time are constrained by the knowledge and technological resources of that time. These systemic variables have led to countless historical cases of major scientific breakthroughs being achieved nearly simultaneously by competing inventors or scientists. Another example is the progression of civil rights. Once certain rights are attained, the battlefront moves towards other unsecured rights. In these examples, the specific actors

that move society forward seem less critical. The future behavior of a group appears predetermined by the goals of the group as a whole and the constraints placed on it.

The question of how people address the challenges associated with achieving coherent group action requires an answer that spans multiple levels of analysis. This question is central to the "micro-macro divide". Yet much like how scientists broadly do not understand the ways in which cognitive abilities emerge from neuronal activity, we lack general solutions to this "micro-macro problem" in understanding social systems.

While the micro-macro problem is widely considered to be unsolved by sociologists [98], particular paradigms do offer partial solutions. In economics, many macro-level theories have micro-level foundations [75]. In the case of these economic theories, macro-level outcomes provably emerge from combinations of certain micro-level individual behaviors, although the micro-level assumptions are often untested or empirically invalid [17]. Analytical sociology takes a different approach [48]. Analytical sociologists, as with mathematical and computational sociologists before them, are more inclined to study aggregate patterns in simulations of agent-based models and the like. These scholars propose simple heuristic rules, some of which have empirical support, and study the behavior of simulated social systems composed of agents using these rules. This approach trades the detailed analytical understanding economists achieve for a coarser understanding that is grounded in more plausible micro-level behavioral models.

Computation offers another solution, underappreciated in many of the social sciences. Computer scientists constantly deal with differing levels of analysis. Computational problems are solved by algorithms. Algorithms are written in high-level programming languages. High-level programming languages are implemented by low-level programming languages, which themselves ground out in bit operations in a computer's machinery. Cognitive scientists recognized that the same sorts of reasoning computer scientists use to understand these multi-scale computer systems can apply to any information processing system [90, 79].

1.1.1 Levels of Analysis

My application of computational thinking to understanding social systems builds directly on the Marrian framework of analysis from the cognitive sciences [79]. Marr proposed that any information processing system could be understood at three levels of analysis: one that describes the computational problem the system is solving, one that describes the algorithms and representations being used to solve that problem, and one that describes how those algorithms and representations are implemented. I attempt to view groups of people as distributed information processing systems and apply a similar set of levels of analysis. However, when conceptualizing groups as distributed information processing systems, an additional complexity arises, which is the connection between the individual algorithms and the computation of the group as a whole. Concepts from the theory of distributed algorithms assist in resolving this complication.

Computational Level

The first level of analysis I consider is directly analogous to Marr's: the computational level. In a group context, the computational level of analysis consists of the computational problem that the group is facing. The contribution of the first part of this thesis is to characterize generic computational problems that groups face in attempting to achieve effective goal-directed action, and subsequent chapters study how these problems are solved. For example, a significant part of this thesis addresses the question of how groups arrive at accurate shared beliefs. One computational problem associated with accurate shared belief formation is information aggregation. Establishing that groups do aggregate information is an accomplishment at the computational level of analysis. Information aggregation is a part of the computation that groups perform.

Group Algorithm Level

The second level of analysis that Marr considered was the algorithmic level of analysis. This level consists of understanding how the computations an information process-

ing system performs are achieved. I break this level of analysis into two, because in certain distributed computational systems there is a simple way to characterize what algorithm is being used as well as a more complex way to describe the same algorithm. That is, it is sometimes possible to understand a distributed algorithm in terms of an implicit centralized algorithm or implicit centralized representation that is effectively being implemented. This "group algorithm" provides the basis for interpretable macro-level models of group behavior. For example, in transactive memory systems [124], knowledge is represented in a distributed manner across individuals, but we can reason about the group as a whole via what memories are accessible through the group members' interactions.

Individual Algorithm Level

In addition to the level of the group algorithm, I also will study the individual level algorithm. This level of analysis describes the algorithms and representations that individual group members are using to achieve either the group algorithm or the computation of the group. This level of analysis is the most directly akin to Marr's algorithmic level.

Heuristic Implementation Level

Finally, I also consider how individual algorithms can be achieved through heuristic implementations. For instance, I will study a case where people act in a way that is consistent with Bayesian updating but in which this updating can be implemented by a simple non-Bayesian heuristic behavioral rule. This level of analysis relates to intermediate levels of analysis proposed recently in cognitive science [35, 42].

1.1.2 Example: Waiting in Line

A simple example that illustrates these four levels of analysis is waiting in line. A line at the deli implements the computation of a first-in, first-out (FIFO) queue. The group algorithm that is used to solve the FIFO queue problem is to maintain a linked

list data structure between elements of the queue, and pop elements off the list as needed. The individual algorithm that implements this linked list is for each element of the list (each person in the line) to keep track of who is ahead of them in the list. The heuristic implementation used to keep track of who is ahead of you in the line is simply to stand behind that person.

1.2 The Organization of this Thesis

I leverage this computational approach to understanding social information processing systems in order to develop a general framework for understanding certain aspects of human social and collective behavior. I will begin by presenting the overarching mathematical framework for characterizing collective intelligence that I use to organize my own thinking about social systems. A simple result from multiagent system suggests that coherent group behavior and collective intelligence in the form of rational collective agency arises from three conditions: aligned utilities, beliefs, and actions. This framework then serves as an informal point of inspiration to motivate my modeling efforts. For the remainder of this thesis, I focus predominantly on the second condition of aligned beliefs. I study how people align their beliefs in processes of shared belief formation. I first present an illustration of the emergence of collective agency from an online multi-participant laboratory experiment. In this example, groups achieve better performance than individuals via implicit communication leading to more accurate shared beliefs about their environment. I then investigate how this kind of implicit communication mechanism that people use to learn from each other manifests in a complex real-world example. Finally, I examine how people arrive at accurate interpersonal beliefs about what actions to expect from each other in order to achieve coordinated actions. I conclude with a discussion of further directions in this research program.

Chapter 2

Three Conditions for Collective Rationality

When the actions of multiple agents mesh well together, a group can accomplish much more than the individuals in that group alone. We have built the most complex society the earth has ever seen together. Cumulative scientific discovery continues to allow us to accomplish previously unimaginable feats of engineering. We have collectively refined our cultural sensibilities to distinguish fine differences within a range of arts. Governments are more or less largely functional. Companies maintain corporate identities. Community organizers bring people together for action on all scales. These collective accomplishments motivate the study of how groups solve problems collectively, much as the remarkable problem-solving abilities of individual humans motivates the study of cognition.

And indeed, just as we can think about individual cognition as computation or information processing, we can think about certain types of collective behavior as a distributed computation. This idea is referred to as "distributed cognition" [55]. The concept of distributed cognition has the potential to allow us to understand puzzling phenomena in group behavior, such as the nature of human collective intelligence, the emergence of structure in society, and cognitive regularities in culture. However, a purely analogical comparison to individual cognition or computation fails—there are fundamental differences between how individual computational agents operate

compared to how distributed systems operate. Distributed systems often do not have access to physical synchronization mechanisms or shared memory, and may not even be implementing the same plans.

How do we negotiate our disparate and incommensurable personal worlds to achieve coherent group behavior? What leads to coherent information processing and group action from such a mess of individual behaviors? The key to answering these questions lies in our scientific understanding of agency. Coherent groups are often modeled or conceptualized as agents. In economics, a prominent model is to assume firms act rationally (e.g., [57]). In political science, nation states are primary actors (e.g., [105]). Organization science studies organizational learning and memory (e.g., [73, 78, 122]).

Rational choice theory is a powerful theory of action that allows us to reason about coherent agential behavior [11]. A rational agent is endowed with beliefs about the world and preferences or goals, specified as a utility function. Rational agents act to maximize their expected utility. Viewing coherent group behavior through the lens of rational choice theory, the question of how coherent group action arises becomes: when do groups of rational individual agents achieve rational collective agency?

In the present chapter I explore the implications of a basic mathematical result from the literature on multiagent systems that offers a simple potential answer to this question: three intuitive conditions suffice within the rational agent modeling framework to guarantee rational collective agency—aligned utilities, beliefs, and actions. Each of these conditions is also associated with impossibility results from social choice theory, philosophy, and computation. The impossibility results illustrate that achieving these conditions is not easy, which creates a barrier for accomplishing group agency. These barriers in combination with the potential benefits of acting together in groups create pressure for mechanisms that help to achieve the conditions that lead to rational collective agency. My central proposal is therefore that certain aspects of human social behavior and human society have been biologically or culturally evolved to function, at least in part, to achieve the conditions for rational collective agency—alignment in utilities, beliefs, and actions. I use this proposal as a motivation to

study mechanisms that might promote each of these conditions.

To support this proposition, I survey several well-established empirical findings about human social behavior from social psychology, sociology, anthropology, economics, and political science. I show how these findings can be understood through the lens of collective agency. The perspective I introduce unifies several empirical regularities that have been observed in human social interaction and society. We can view phenomena as varied as homophily, social influence, social learning, and conventions as promoting rational collective agency by making groups more homogeneous. I also argue that many of these phenomena are enabled by the power of human theory of mind—one of the attributes of humans that is thought to possibly distinguish us from other animals [115]. My work brings a mathematical cognitive perspective to bear on the basic questions of understanding collective intelligence. My primary aim is to build a bridge between researchers in computational cognitive science, distributed cognition, collective intelligence, and computational social science. A central utility of my perspective is as a source of inspiration for research questions, and a source of clarity in reasoning about collective behavior.

The basis of rational collective agency: aligned utilities, beliefs, and actions

The theoretical result I build upon is that rational collective agency arises from aligned utilities, beliefs, and actions among individual agents. I offer a formal statement and an informal proof sketch of this proposition before discussing its implications.

Problem Statement

I assume the setting of a finite time horizon decentralized multiagent partially observable Markov decision process (Dec-POMDP) [12]. Such a process consists of a fixed set of N agents, discrete time, a finite time horizon T, a set of states S, a set of actions for each agent A, an observation distribution $P(o_i \mid s)$ for each agent i, and a

stochastic transition function $P(s' \mid s, \boldsymbol{a})$. I also assume the group has a shared goal that is represented as a joint reward function $R : S \times A^N \to \mathbb{R}$. This final assumption restricts the theory to settings in which the group has a well-defined collaborative objective. However, the individual agents also have their own personal reward functions, $R_i : S \times A^N \to \mathbb{R}$.

The best a group can achieve in this situation is to take the joint actions that maximize its the expected cumulative reward at each step, according to all of the observations all group members have received up to that time:

$$\{\boldsymbol{a}_t^*\} = \underset{\boldsymbol{a}_t}{\operatorname{argmax}} E[\sum_{t'=t}^{T} R(s_{t'}, \boldsymbol{a}_{t'}) \mid \boldsymbol{o}_{\leq t,\cdot}].$$

I say that rational collective agency is achieved when the group of agents takes a joint action in this optimal set at every time step. Rational collective agency therefore corresponds to executing an optimal *centralized* multiagent POMDP policy. The challenge in executing such a policy is that the agents are decentralized. In particular, *a priori* the agents only have access to their own noisy state observations and incomplete knowledge of what actions other agents will choose on a particular step. What abilities must a group possess to achieve this ideal?

It is straightforward to show that aligned utilities, beliefs, and actions are necessary and sufficient attributes for rational individual agents to have in order to achieve rational collective agency. More precisely, a group can implement an optimal centralized policy, for an arbitrary Dec-POMDP, if and only if (1) each agent ignores its personal reward function, acting only to optimize the group reward function, R; (2) all agents have accurate shared beliefs, i.e. each agent believes the state of the world is s with probability $P(s \mid \boldsymbol{o}_{\leq t,\cdot})$ at each time t; and (3) agents have a coordination mechanism that allows them to choose a unique joint action among the set $\{\boldsymbol{a}_t^*\}$.

Proof Sketch

I now present an informal argument for why aligned utilities, beliefs, and actions are sufficient for rational collective agency. Since the agents have accurate shared beliefs,

they can all compute the set of optimal actions, $\{\boldsymbol{a}_t^*\}$. Since the agents ignore their personal reward functions and act only to optimize R, all agents will choose an action consistent with a joint optimal action if they can. The assumption that the agents have access to a coordination mechanism ensures that a unique joint action can be selected for the group.

We can also show that these conditions are necessary for rational collective agency. If any one of the agents' utilities, beliefs, or actions are unaligned, there will exist a Dec-POMDP in which the group will fail to achieve rational collective agency. We can prove this converse statement by providing examples of problems where the group fails in each case. When agents act according to their own personal reward functions instead of the group reward function, we only need to postulate a one-step Dec-POMDP in which the group reward conflicts with agents' personal rewards. When agents do not have accurate aligned beliefs, we need only construct a one-step Dec-POMDP in which agents report their beliefs, and the group reward is a local proper scoring rule that incentivizes honest reporting. In this case, the group utility is maximized by all the agents reporting what would be the optimal shared beliefs—which the agents do not know and hence cannot achieve. Finally, when agents do not have a coordination mechanism, they must either attempt choose probabilistically among the set of optimal joint actions, or choose a suboptimal joint action. Probabilistic attempts can lead to coordination failure, however. A simple N-agent pure coordination game illustrates the difficulty. In such a game the agents must each choose one of two actions, but must all choose the same action in order to receive a positive state-independent group reward. Without access to a coordination mechanism, the agents can achieve at best a probability 0.5^N of positive reward, which is clearly a lower expected reward than the optimal joint action given the ability to coordinate.

Example

To illustrate that mechanisms could exist that meet these conditions, I now present an informal example illustrating how the three conditions of aligned utilities, beliefs, and actions could be achieved in a specific case. Consider a group of friends who want to go

out to dinner together. Different friends might have different preferences, but suppose that the group agrees to try to go to the best burger restaurant in the neighborhood. The friends then have an aligned utility function on the quality of the burgers they will eat. If the friends look at all the reviews they can find of each burger joint, and also share all of their own personal experiences with each restaurant, then the friends will have aligned shared beliefs about the quality of each restaurant. There may still be multiple places that seem equally good after this process of information sharing. To coordinate on which location to choose, the friends could agree on a tie-breaking mechanism, such as going to the best-looking restaurant that is geographically closest to the mean of the friends' locations, or letting the youngest member of the group decide.

A more complex example, which will be explored in more detail in a later chapter of this thesis, is an online community that seeks to identify good content. In this case the shared goal of the group can be viewed as identifying and consuming the best possible content. One way people can effectively aggregate beliefs about content quality is to probabilistically sample recommendations from peers in the community, then assess those recommendations personally. This social sampling mechanism leads to content popularity reflecting accurate beliefs about content quality [69], which thereby allows people to arrive at accurate shared beliefs. This social sampling mechanism also allows individuals to have an optimal probability of consuming good content over time, since random peers are likely to have good recommendations. This form of peer recommendation therefore leads individuals to act with the benefit of all of the accumulated knowledge the community has at its disposal.

Discussion

These examples illustrate weak forms of rational collective agency—rational collective agency in specific cases. However the definition of rational collective agency I give—optimal performance across all tasks—is quite strong, and is most applicable to groups that face a variety of different tasks. Furthermore, certain particular tasks will not require all three conditions be satisfied. A group therefore may not need all

three conditions to be satisfied at all times, or potentially at any time, depending on what tasks are faced by the group. The conditions themselves are also quite strict. The result states that if agents have exactly equal beliefs and exactly equal utility functions, they can essentially act like a single agent. This nearly tautological result is nevertheless interesting as a point of inspiration. Mechanisms for achieving "strong", problem-independent rational collective agency may or may not exist. The impossibility results I will review shortly suggest that strong rational collective agency may only be approached or approximated. We will nevertheless see that human social behavior and society display a collection of mechanisms that may function to raise the level of collective agency. Having exactly equal beliefs and utilities among people in a group is unreasonable to hope to achieve. Future theoretical investigations could attempt to determine when and to what extent approximating the alignment conditions leads to approximately rational collective agency.

How the framework of rational collective agency relates to social systems

We see that mathematically it is straightforward to show that ideal rational group agency depends on the three conditions of aligned utilities, beliefs, and actions. However, these mathematics rely on a foundation of pure rational choice that we know to be empirically questionable. My goal here is to argue that this mathematical result still provides insight into the structure and function of social systems. Given that utility function alignment, information aggregation, and coordination pose fundamental computational problems, we can think about what aspects of social systems might function to help alleviate these problems in effective groups. In particular, each of these conditions is associated with certain impossibility results. A foremost difficulty in achieving the conditions for rational collective agency is therefore circumventing the impossibility results associated with each condition.

	Example
Aligning Utilities	Homophily
	Social influence
	Social preferences
	Joint intentions
Aligning Beliefs	Social learning
	Opinion polls
	News/social/public media
Aligning Actions	Social norms
	Law and conventions

Table 2.1: Examples of social phenomena that promote the conditions for rational collective agency. See text for details of each example.

Examples of how existing empirical results relate to this framework

Social processes that align preferences

The first condition for rational collective agency is aligned utility function. In order for groups to avoid conflicting objectives, the utility functions that people act on must be aligned. A well-known impossibility result from social choice theory, known as Arrow's Impossibility Theorem [6], states that there is no general mechanism for fairly aggregating arbitrary preferences of group members. Taken at face value, this result suggests people may never be able to attain rational collective agency without resorting to centralized authority because there will always be cases where a fair mechanism cannot be identified to resolve differences between individual utility functions. However, natural mechanisms appear to exist for both group formation and utility function adjustment that could circumvent this theoretical impossibility result.

One of the most widely observed empirical regularities in the study of human social networks is the phenomena of homophily [83]—the tendency for people who are similar to each other to tend to interact preferentially. Homophily has been observed along a range of attributes, from various demographic characteristics [83] to personality traits [128]. If homophily occurs along the dimensions of people's utility functions, then this ubiquitous social process would naturally lead to groups that have

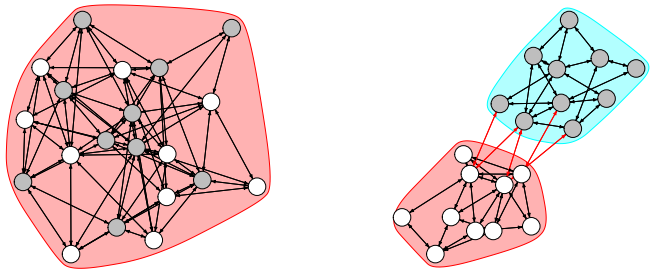

Figure 2-1: **My framework suggests viewing the widely observed social process of homophily as functioning to help align utility functions.** This figure displays two networks of nodes representing people that have preferences for one of two outcomes. Node preferences are shown as the color of the plotted nodes. Results of community detection illustrating plausible group splits according to network edges are also shown. (Left) A network with attributes distributed randomly among nodes. (Right) A network displaying homophily along the plotted node attribute.

better aligned utility functions. My theoretical results therefore suggest a functional lens for interpreting this well-known social phenomenon. Homophily may help to circumvent the impossibility of universal social choice mechanisms for agreement on preferences.

In addition to homophily, social influence on preferences is another widely observed phenomenon [25]. When people are not self-selecting social partners based on attribute similarity, they influence each others' preferences directly. Again, social influence has been shown to occur across a wide range of types of preferences, from exercise habits [4] to online content [89]. Social influence and homophily together may provide a powerful force to align utility functions.

Another hypothesized feature of human behavior that would lead to better aligned utilities is social preferences. Social preferences are a postulated modification to individual utility calculus that incorporates the utility functions of other actors into one's personal utility function [36]. Social preferences have some empirical support, and are advocated as a solution to puzzles in observed human behavior, such as excess levels of cooperation as compared to the predictions of pure individual rational action [97].

A related but distinct thread of literature hypothesizes that people are able to adopt shared group utility functions via joint intentions. These joint intentions involve group members reasoning about what would be best for everyone if they agreed to act as a group, and thereby allows the group to attain Pareto optimal actions. This logic appears in economics under the label "team reasoning" [109] and in philosophy and psychology as joint or shared intentions [16, 116]. These models have recently been empirically tested in psychology experiments, establishing that people may indeed participate in a kind of virtual bargaining that leads to fair shared utility functions being adopted by individuals in group contexts [63].

The underlying cognitive capabilities that enable these various social processes is likely to be the powerful inference mechanisms that allow us to infer each others' intentions and preferences. These cognitive abilities, part of our theory of mind, are likely to underlie the processes of homophily, social influence, social utility, and joint intentions.

Social processes that align beliefs

The second condition for rational collective agency is accurate aligned beliefs. In order for groups to make accurate decisions, information must be aggregated. As with preference aggregation, there are impossibility results about belief aggregation. In particular, the Discursive Dilemma shows how majority vote on beliefs can lead to inconsistent sets of aggregate beliefs [74]. However, the picture here is somewhat more optimistic, at least as long as people update their beliefs rationally, and information is plentiful. Aumann's agreement theorem states agents will always come to agreement if they have common priors [8]. Elementary asymptotic probability shows that if agents have unlimited information and identical mental models of the world, they will also eventually agree, even if their specific priors differ. At the same time, people may easily differ in their priors or mental models of the world, or lack infinite information about a situation.

Several features of human cognition support belief agreement processes. Social learning—the ability of people to learn from observing each other—is widespread

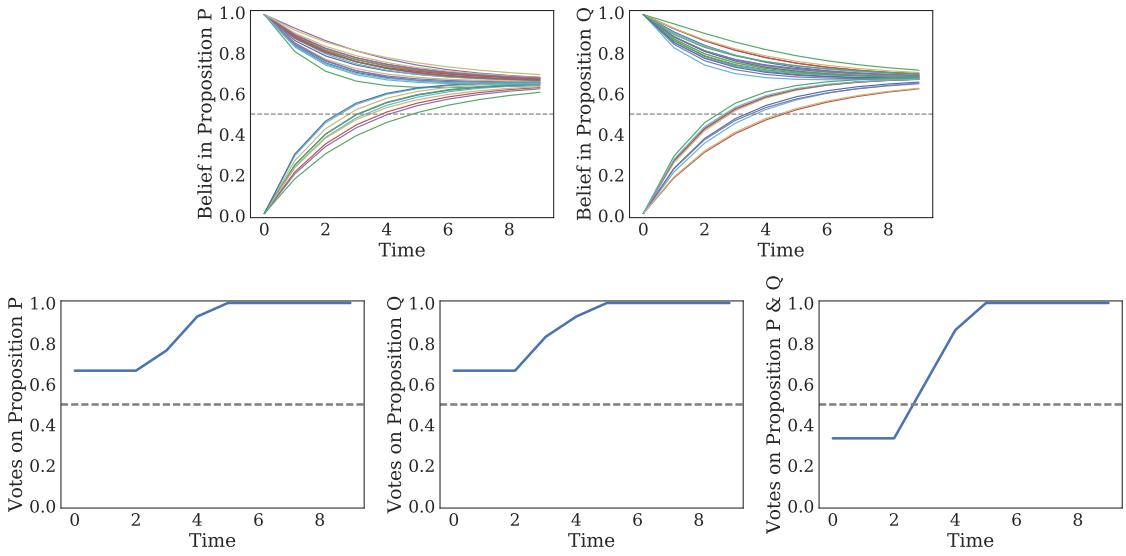

Figure 2-2: **Social learning helps resolve aggregate inconsistencies in group judgements.** The Discursive Dilemma describes a paradoxical situation in which the truth value of three propositions—P, Q, and $P\&Q$—are decided based on majority vote. Beliefs about P and Q can be distributed amongst the population in such a way that a majority of the population believes P, and a majority believes Q, but only a minority believe $P\&Q$. I simulated this situation to see if individuals influencing each others beliefs in a random network would resolve this dilemma. At each time step, individuals update their beliefs according to the Degroot learning rule [29], which is a common model of heuristic social learning. (Top) Beliefs of each individual over time. (Bottom) Aggregate votes over time. The horizontal dotted lines indicate the majority-rule decision boundary.

and has been hypothesized to underlie human society's remarkable ability at cultural accumulation [15]. There is still debate about whether people engage in social learning through rational mechanisms, heuristic mechanisms, or semi-rational mechanisms [106]. Regardless, these models all share the characteristic of bringing peoples' beliefs to be closer to one another.

In addition to learning by obtaining information from others, people also have a tendency to align their mental models of the world through exposure with each other. Philosophers have argued that the construction and maintenance of the ontologies that we use to understand the world are socially determined and reinforced [10]. Aspects of management and organizational training focus on this process of onboarding by aligning mental models [86]. Presumably, once people's mental models and priors

are aligned, then social learning can lead to agreement on specific beliefs as needed.

Once again, these social mechanisms are all likely supported by complex cognitive abilities. Just as our theory of mind allows us to infer others' preferences, so does it allows us to infer others' beliefs. Humans develop an ability to reason about others' beliefs at an early age. These mind reading abilities undoubtedly facilitate a range of our social learning mechanisms. Models of social learning—whether rational or heuristic—invariably assume exposure to and knowledge of others' beliefs, but beliefs per se are never directly observable—they are abstractions that must be inferred.

We as a society have also established aggregation mechanisms that facilitate shared belief formation. Opinion polls are a mechanism for information aggregation. The news media and the entire publishing industry serves to communicate accumulated knowledge. Modern developments such as social media and social recommendation systems also serve the function of aligning beliefs. Recent research has shown that these systems are remarkably efficient at aggregating information [20], suggesting that an implicit force favoring accurate shared belief formation may be helping to shape the design of these platforms.

Social processes that align actions

The final condition for rational collective agency is aligned actions. Individuals must be able to select actions that mesh well together. The group must be able to select one joint action from the set of optimal joint actions. Certain components of an optimal joint action may be risky in the sense that an individual's action can be bad if others do something unexpected, and hence the individual taking a risky action must be supported by the rest of the group in that joint action. Once again, there are impossibility results associated with perfect coordination. Computer scientists have shown that exact coordination in asynchronous systems cannot be guaranteed [45, 76]. Fortunately, although perfect coordination is impossible, rational coordination is theoretically achievable. Rational coordination can occur when the risk of attempting to coordinate is balanced by the utility gained by successful coordination [87]. However, even rational coordination appears to require impossibly complex mental representa-

tions. Rational coordination requires infinitely recursive interpersonal beliefs similar to the logical concept of "common knowledge"—I know that you know that I know, etc. that you will act how I expect you to act.

As in aligning utilities and beliefs, social influence may play a role in aligning actions. Classic studies display human susceptibility to social norms [7, 85], a tendency which has also been confirmed in large-scale observational studies [81]. Moreover, recent computational work has shown that norms and norm learning are feasible ways to achieve multiagent coordination [52].

Another possibility is that people use tractable finite representations of common knowledge [65]. Such representations could support richly structured joint intentions. Joint intentions are thought to underlie many types of coordination behavior and have also been hypothesized as another factor in what makes human unique among animals in our complex society.

Laws and societal conventions also clearly play an important role in solving many coordination problems. A classic example is deciding on which side of the road cars should drive. The perspective that I highlight here is that these laws and conventions might not just assist in coordination, but may also facilitate the more general emergence of rational collective agency, for example in achieving a distributed solution to the computational problem of routing inherent in traffic.

Collective agency and previous accounts of "group mind"

The framework I propose extends a long line of research in sociology, social psychology, and organization science. Early theories of collective agency once played an important role in debates about the workings of society, but were largely discredited in sociological discourse. The general character of these theories was that groups had some inherent agential properties, or at least causal force, which would influence how the individuals in those groups behaved. For instance, Durkheim speaks about "social facts" that exist only in the collective consciousness, not in any particular individual, but nonetheless exert causal forces on individual action [30]. Some twenty

years after Durkheim's writings, this type of perspective was forcefully criticized by Allport, who argued that there is no inherent explanatory power in macro-level descriptions of social processes [2]. Allport's perspective continues to be dogma among most sociologists today.

My approach to understanding collective agency does not require that groups have any inherent agency or causal power above and beyond that of the constituent group members. I simply seek to characterize when groups as a whole act in accordance with the constraints of rational collective action. The tack I take is still not without controversy. A complex systems or network perspective of groups and society suggests that people will never fit so neatly together as to form perfectly coherent groups, just as people individually rarely conform perfectly to the predictions of rational choice theory. Other complications include defining group boundaries and demarcating what actions of an individual are taken in that individual's capacity as a group member.

While collective agency may still be a simplified representation of the complexities of collective behavior, there are certain areas of social life that clearly rely on the formation and maintenance of approximately coherent groups. Organizations, teams, and group problem solving are all prominent examples. Even Allport conceded that certain features of collectives, such as cultural phenomena, had coherent macro-level process descriptions. Indeed, recent authors in the literature on human collective intelligence have shown empirically that there are measurable attributes of human team ability that are distinct from individual member ability [126]. Together these facts suggest that coherent group agency may indeed be a useful concept for reasoning about at least some facets of human collective behavior.

Beyond the classical literature relating to these topics, several computational theories of coherent group behavior besides my own have been previously proposed that ground the dynamics of collective action in micro-level foundations. One early framework was coordination theory [77]. Coordination theory provides abstract representations of complex coordination problems, as well as components that are useful in solving those coordination problems.

A notable contribution from the computer science literature is the SharedPlans

framework [44]. SharedPlans is a normative framework that describes a set of sufficient, computationally tractable conditions for collaborative activity. The focus of SharedPlans is the complex web of interpersonal intentions that leads individual plans to mesh into coherent group behavior. SharedPlans is not explicitly related to collective agency, but could serve as a representation for complex joint intentions, which aid in alignment of utilities and actions.

The deeply related topic of distributed cognition has also been offered recently as a lens for understanding an array of social and collective behaviors. Distributed cognition occurs when groups of people implicitly perform computations together, perhaps assisted by artifacts and tools in their environment. The word "cognition" is used because the computation that group are performing is analogous to the computation that is viewed as underlying individual cognitive processes. Hutchins, who coined the term, studied what is now a classic example of the crew of a navy ship solving the computational problem of navigation on the open waters [55]. The idea that Hutchins proposed was that we could think of the crew, in combination with the tools and procedures in their environment, as unified information processing system. Viewing groups as information processing systems allows them to be amenable to Marrian multiscale analysis [79]. Others have taken up Hutchins' program in the context of modern digital systems [53].

A similar perspective appears in a prominent thread of work within social psychology, the study of transactive memory [124]. The idea of transactive memory is that people can enhance their memory as a group by engaging in behaviors such as selective distributed storage and retrieval. People as a group then appear to have a cognitive ability analogous to an individual cognitive ability, but with higher capacity, implemented at the group level.

A distributed cognition perspective also appears in the area of "social computing" within computer science [23]. Social computing is sometimes used literally to mean crowdsourcing, where a curated group of people explicitly execute a distributed computational task together. The term is also sometimes used to describe collective behavior within sociotechnical systems such as social recommendation systems, in

which the community of users of a platform collective discover or curate information.

The usefulness of the concept of distributed cognition stems from its relationship to distributed computation, in combination with the computational understanding of the mind. Our modern understanding is therefore that there is not a "group mind" that operates with a kind of downward causality controlling what individual people do, but that groups can execute cognitive processes distinct from individual cognitive processes via distributed computation implemented by the combination of individual actions. Yet, despite these theoretical advances, there is minimal formal basis outside of philosophy for these ideas of "distributed cognition", "social computing", or "collective intelligence". My contribution is to lay out a general formal framework for collective agency, and to show we can derive new insights into the nature of human behavior and the structure of human society from this framework.

Concluding remarks and future directions

The formal lens of rational collective agency provides a broad and concise framework for reasoning about collective behavior and collective intelligence. I have argued that this framework (1) yields a formal way to understand the concept of a "group mind" without sacrificing reductionism, and (2) suggests a macro-level functional interpretation of disparate social phenomena. My framework can also serve as a source of inspiration for research questions. What aspects of society promote collective agency? How do individuals leverage or manipulate collective agency in order to enrich their own power? Can aspects of strategy in warfare, crowd control, or espionage be understood as undermining conditions for collective agency in opposing groups? Can we promote effective collective behavior by designing mechanisms to support the conditions for rational collective agency?

The fact that we can show that the three conditions of aligned utilities, beliefs, and actions are both necessary and sufficient for collective agency also shows that, at least within the rational agent account of human behavior, groups need not worry about other factors besides these three to achieve rational collective agency. Aligned

utilities, beliefs, and actions are "all a group needs". Of course, other factors also seem to be important to group behavior, such as individual motivation, team culture, and emotional factors. Although my current framework precludes the primal importance of these factors for establishing effective groups, perhaps richer theories of individual agency that better incorporate these factors could also be extended to enrich my framework of collective agency.

What forces shape human social behavior and the structure of human society? Here I have identified three endogenous forces—attributes that, if achieved, can transform a group of isolated individual into a rational collective that can achieve much more than the individuals alone. These attributes themselves require underlying cognitive capacities. Although there is controversy around the arguments (e.g., [71]), two of these capacities—human theory of mind and joint intention—have been offered as traits that are in fact distinctly human, distinguishing us from other animal species [115]. However, we have also seen that people are endowed with a number of other behavioral mechanisms, such as various forms of social influence, that can be viewed as promoting rational collective agency. Perhaps not just joint intentions, but the unique combination of flexible solutions to all three conditions for rational collective agency is part of what makes human unique. Perhaps part of what makes humans special is our ability to create coherent collective agency on-the-fly with unrelated individuals in a variety of contexts due to the solutions we have for achieving the conditions for rational collective agency.

Chapter 3

Observing The Emergence of Group Agency

In this chapter I present an example that illustrates the emergence of collective agency in an experimental setting [67]. I focus on analyzing the behavioral mechanism that leads to accurate shared belief formation in this context. Despite its importance, human collective intelligence remains enigmatic. We know what features are predictive of collective intelligence in human groups, but we do not understand the specific mechanisms that lead to the emergence of this distributed information processing ability. In contrast, there is a well-developed literature of experiments that have exposed the mechanisms of collective intelligence in nonhuman animal species. I adapt a recent experiment designed to study collective sensing in groups of fish in order to better understand the mechanisms that may underly the emergence of collective intelligence in human groups. I find that humans in this experiments act at a high level like fish but with two additional behaviors: independent exploration and targeted copying. These distinctively human activities may partially explain the emergence of collective sensing in this task environment at group sizes and on times scales orders of magnitudes smaller than were observed in fish.

3.1 Introduction

Many common examples of collective behavior illustrate apparent failures of collective intelligence. Mobs, market panics, and mass hysteria draw attention because of their perceived irrationality and drastic consequences. However, the successes of collective intelligence are as remarkable as the failures are devastating. The richness of human culture, the incredible pace of our technological developments, and the gradual progression of our scientific understanding of the universe stand out as both distinctively human and heavily reliant on the emergent behavior of the interactions of many individuals. Even at a less grandiose level, humans regularly agree to work together to accomplish tasks that no individual could accomplish alone via dynamic cooperative interactions that are hypothesized to be uniquely human [115]. Yet little is known about the specific mechanisms underlying these synergistic processes of self-organization.

Many mathematical and computational of models collective behavior have been proposed. However, as a result of the logistical difficulties in conducting real-time human experiments involving multiple participants, and as a result of a broader lack of data analysis aimed at understanding collective behavior, the quantitative study of collective behavior has largely lacked an empirical basis. Recently, researchers have begun conducting carefully controlled laboratory experiments to test and refine models of collective behavior [27, 38]. Yet many of these experiments, with some notable exceptions [39, 60], have been conducted using nonhuman animal subjects. We are therefore quickly developing a better understanding of the collective behavior of ants [95], bees [103], cockroaches [3], and fish [123], but our empirically-grounded quantitative understanding of human collective behavior remains limited.

In the present chapter I harness recent technical advances in running real-time, networked experiments on the web [47] to develop and test a model of collective human behavior. I build on a recent experiment designed to study the collective behavior of a particular species of fish [9] that is one of the clearest illustrations of collective intelligence in a nonhuman animal group. In this previous experiment, the researchers

studied a type of fish called the golden shiner that prefers to spend time in dark areas of the water, presumably to avoid predators. Aware of this natural propensity of the fish, the researchers projected time-varying spatially correlated light fields into a fish tank. The researchers then studied the effectiveness of the fish at finding the darker areas of the tank as a function of the number of fish participating in the task. The researchers found that average group performance increased significantly as a function of group size, and they identified two simple behavioral mechanisms driving this improvement: First, individual fish tended to move more slowly in darker areas. Second, individual fish also tended to turn towards conspecifics. The researchers argued that the combination of these mechanisms generated an emergent collective gradient sensing ability in groups of fish that had been absent in individual fish.

This experiment provides a beautiful example of a higher level of intelligence at the group level emerging from minimal intelligence at the individual level. However, while these simple mechanisms did appear to give rise to surprisingly effective group behavior, they only lead to substantial gains in performance for large groups of 50 or more fish. In contrast, we expect humans in a similar task to show significant gains with much smaller group sizes. In particular, we expect that humans should be able to make use of theory of mind, an ability to draw inferences about the underlying mental states of other players, to better utilize social information in a similar environment.

To elucidate these potential differences between humans and fish, I developed a version of the gradient-sensing task for human participants. Specifically, I recreated the environment used by [9] as an online real-time multi-player game. In this experiment, participants controlled avatars in a virtual world. Every location in this world corresponded to a score value that changed over time, and participants were awarded bonuses proportional to their cumulative scores in the game. The score of a player at a particular point in time was simply determined by the location of that player in the virtual world. The incentives for participants to achieve high scores were designed to parallel the fishes' preferences for darker areas in their environment. The players either played alone or in groups of varying sizes. I used this virtual environment to investigate how the gradient-tracking performance of human groups changed as

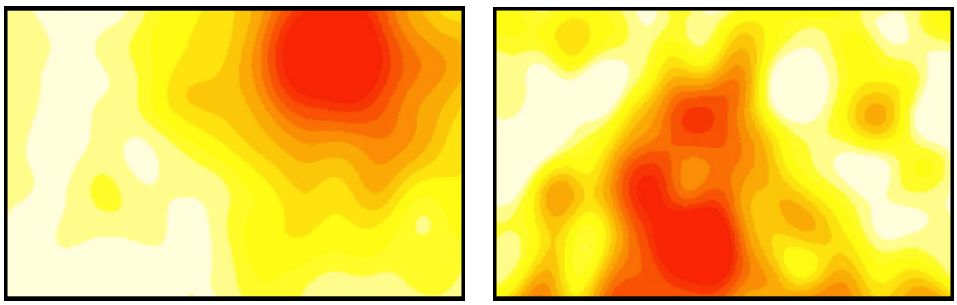

Figure 3-1: Example score fields from the low noise (left) and medium noise (right) conditions at particular points in time. Red areas indicate higher scoring areas.

group size increased, and to attempt to identify behavioral mechanisms underlying collective sensing in human groups.

3.2 Methods

Participants

We recruited 563 unique participants from Amazon Mechanical Turk to participate in our experiment. All participants were from the United States. After excluding 72 participants due to inactivity or latency, and 6 others for disconnecting in the first half of the game, we were left with usable data from 437 participants in 224 groups. These groups ranged in size from one to six individuals. Since we were only able to collect one group of size six, we ignored this group in our analysis.

Stimuli

The game scores of the participants in our experiments were determined by underlying "score fields". These score fields consisted of 480×285 arrays of score values for each 125ms time interval in our game. We generated these score fields using the method reported by Berdahl et al. [9]. First, a "spotlight" of high value was created that moved in straight paths between uniformly randomly chosen locations. This spotlight was then combined with a field of spatially correlated noise. This procedure yields a complex landscape with many transient maxima and a single persistent time-varying global maximum.

We manipulated the weighting between the noise field and the spotlight to generate different task conditions. We used two weight values, corresponding to the "low" and "medium" noise levels reported by Berdahl et al. Examples of score fields are shown in Figure 3-1. 113 individuals (63 groups) were assigned to the low noise condition and 324 individuals (161 groups) were assigned to the medium noise condition. To decrease variability and increase statistical power, we generated only four distinct score fields per noise level, so multiple groups experienced the same fields. To discourage inactivity, players were awarded a score of zero, corresponding to zero bonus, if their avatars were touching a wall.

We attempted to give our participants perceptual and motor capabilities in this environment similar to the capabilities that Berdahl et al. observed in the fish in their experiments. In terms of perception, we restricted the information that participants received about the underlying score fields in the games. We allowed participants to see only the scores at their avatars' locations. The participants could *not* see the scores that other players were obtaining or the scores at any other locations besides their own. However, the positions, directions, and speeds of all other players were visible to each player. All of this information was updated in real-time every eighth of a second. A screenshot of the interface we used for the game is shown in Figure 3-2.

Players controlled their avatars using the left and right arrow keys to turn (at a rate of 40° per second) and could hold the spacebar to accelerate. The avatars automatically moved forward at a constant velocity of 136 pixels per second whenever the spacebar was not depressed. The avatars instantaneously increased to a constant velocity of 456 pixels per second for the duration of time that the spacebar was held down. We chose these speed values to match the speeds that Berdahl et al. reported observing in their fish, and we also matched the playing area dimensions and game duration to the parameters of their experiments. Each participant played in a single continuous game lasting for 6 minutes.

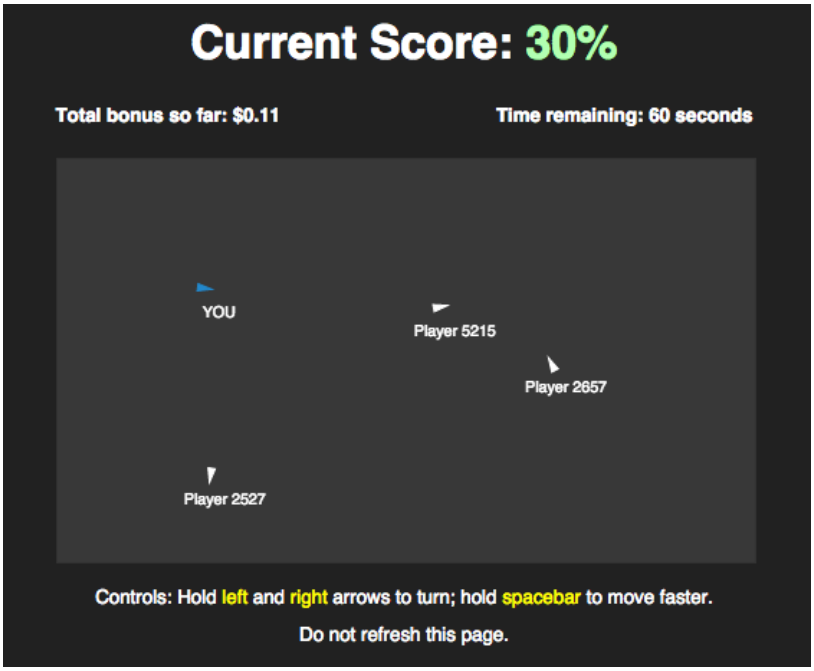

Figure 3-2: A screenshot of the interface that participants saw. The score displayed corresponds to the value of the score field at the location that the player's avatar is occupying.

Procedure

After agreeing to participate in our experiment, participants were presented with a set of instructions. These instructions simply described the mechanics of the game. The participants were not informed about the nature of the underlying score fields and were not encouraged to work together. After successfully completing a comprehension test, participants were then redirected to a waiting room. In the waiting room participants would wait for up to 5 minutes or until a pre-assigned number of other players joined the game. While in the waiting room, participants could familiarize themselves with the controls of the game. Players were not shown any score in the waiting room unless the participant was against a wall, in which case the displayed score would change from a dashed line to a red "0%". We found no evidence for the amount time a player spent in the waiting room having any effect on individual performance in the game (linear regression slope 1.993e-06, with 95% confidence interval [-1e-05, 1.4e-05]). As in the actual game, participants in the waiting room would be removed for inactivity if the player's browser was active in another tab for more

than 15 seconds or if the player's avatar was unmoving against a wall for 30 seconds. We also removed players if their ping response latencies were greater than 125ms for more than 36 seconds. We paid participants 50 cents for reading our instructions, and the participants could receive a bonus of up to $1.25 during the six minutes of gameplay. Final bonuses were computed to be the players' cumulative scores divided by the total length of the game times the total possible bonus. Following the current convention on Mechanical Turk, each participant was also paid 12 cents per minute for any time spent in the waiting room, minus any time that player spent against a wall. These numbers were chosen so that the participants were expected to receive at least the U.S. federal minimum wage of $7.25 per hour for the totality of their time active in the experiment.

We implemented this experiment using the MWERT framework [47]. The MWERT framework uses a set of recent web technologies capable of handling the challenges of real-time, multi-player web experiments, including Node.js, the Socket.io module, and HTML5 canvases. Since MWERT was originally used for two-player games, we had to extend the MWERT framework in several ways to handle the challenges posed by hosting larger groups of players.

3.3 Results

We find that group size is positively related to group performance in this game in the low noise condition. However, we find that there was little effect of group size in the medium noise condition. Average performance as a function of group size in each of these conditions is shown in Figure 3-3. A linear regression on the individuals in the low noise condition produces a significant positive slope of 0.0238 and a 95% confidence interval (CI) of $[0.006, 0.041]$. A linear regression on the individuals in the medium noise condition produces a marginally significant positive slope of 0.0068, 95% CI $[-0.001, 0.015]$, and this trend is weakened substantially with the inclusion of the single 6-person group. Moreover, the marginally significant result in the medium noise condition is driven entirely by the effect of group size in one of the four distinct

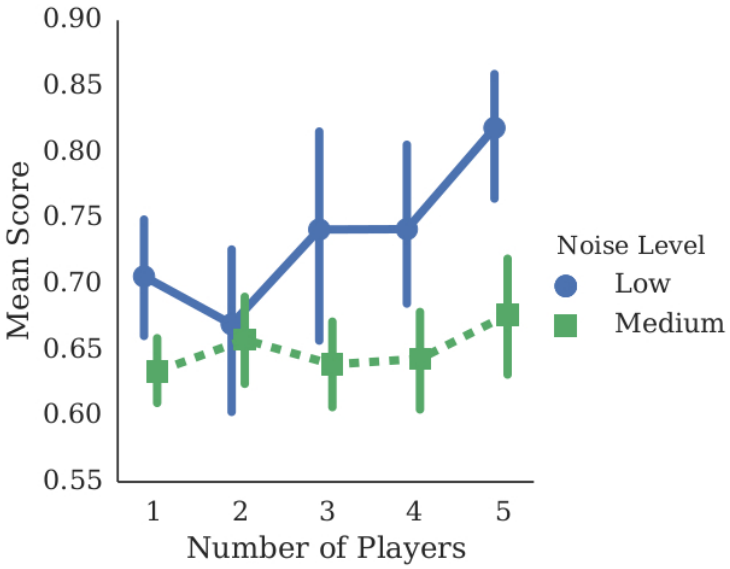

Figure 3-3: Mean performance as a function of group size in the low and medium noise levels. Error bars are 95% bootstrap confidence intervals using the group as the primary bootstrap unit. All points are averages over at least two groups. This plot excludes the single group we were able to collect of size six. Including this group weakens the trend in the medium noise condition.

score fields we used. This particular score field displays a significant effect of group size with a positive slope of 0.0306, 95% CI: $[0.015, 0.046]$, while none of the others do. Qualitative inspection revealed that this particular score field seemed to share spatial properties more similar to the low noise score fields, which may explain the strength of the effect in that particular score field. Overall these results indicate that larger groups do tend to perform systemically better on our task than those in smaller groups, at least in the low noise condition.[1]

In order to understand the factors that may have contributed to the increases in performance achieved by larger groups in the low noise condition, we examine the behavior of the players in our games. We assume a simple state-based representation of player behavior. We then attempt to identify how participants choose to occupy particular behavioral states at each point in time, and we examine the relationship between the players' decisions to occupy particular states and the performance of

[1]Results were similar using a mixed-effects regression including group and score field as random effects, and also revealed larger variability due to score field in the "medium" noise condition than the "low" noise condition.

those players. Specifically, we assume that at any particular point in time a player is either "exploring", "exploiting", or "copying" (cf. [99]). Conceptually, a player is exploring if that player is looking for a good location to exploit, a player is exploiting if that player has found a location where the player wants to remain, and a player is copying if that player is intending to move to the location of another player.

We empirically determine the state of each player at each point in time using a set of hand-tuned filters. All of these filters depend only on information that is observable to any player in the game (i.e., the filters do not depend directly on the scores of any individuals), and hence we can use the inferred states of players as proxies for what other players might infer as the states of those players. Also, since the states are not defined in terms of scores, we can meaningfully quantify the relationship between state and performance.

We now define the three states: exploiting, copying, and exploring. Exploiting a particular location in the environment is not completely trivial for players since the avatars always move at least at a slow constant velocity. In order to attempt to stay in a single location, a player can either meander around a particular location or can persistently hold down one of the arrow keys while moving at a slow speed, which creates a tight circular motion around a particular location. We call this second activity "spinning" because of its distinctive appearance. We then classify a player as exploiting if the player is spinning for 500ms or if the player moves at the slow speed for 3 seconds and has not traveled more than two thirds of the possible distance that the player could have traveled in that time. The second condition is supposed to capture the meandering behavior of individuals who have not discovered how to spin. Copying behavior is more difficult to identify, but appears to often be characterized by fast directed movements towards other players. We thus classify a player as copying if the player is moving in a straight line at the fast speed towards any particular other player consistently for 500ms. We classify a player as moving towards another player if the second player is within 60° on either side of the first player's straight-line trajectory. Finally, we classify a player as exploring if the player is neither exploiting nor copying. Thus a player will be classified as exploring if that

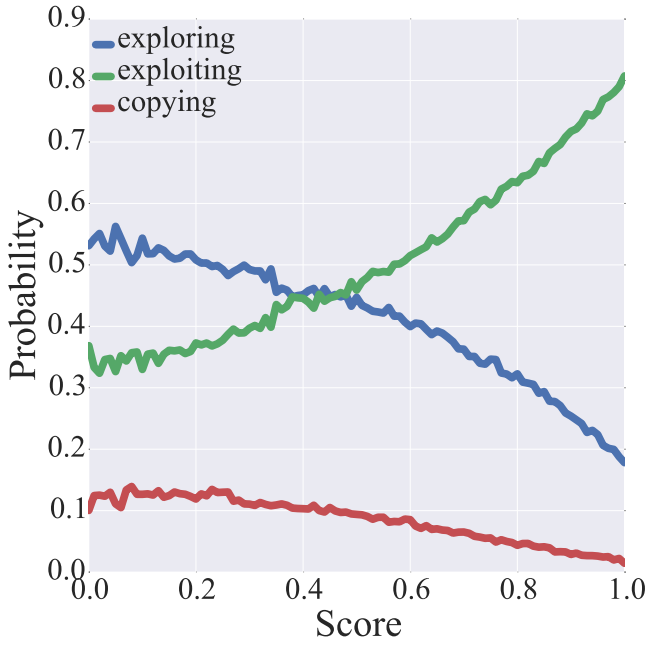

Figure 3-4: The probability of an individual being in a particular behavioral state as a function of the individual's score.

player is either moving slowly but not staying in the same general location, if the player is moving quickly but not towards any particular person, or if the player is moving quickly and turning.

We use these filters to analyze how players behave in our game. First, we compute the probability of a player being in a particular state conditional on the current score that the player is receiving. We find that the probability of a player occupying a particular state is closely related to that player's score. Specifically, players in higher scoring locations are more likely to be exploiting than exploring or copying, but the probability that a player is exploring or copying increases as the player's score decreases. These results, which are visualized in Figure 3-4, suggest that players are choosing their states relatively rationally. Players will tend to remain in good areas and will leave bad areas quickly either by exploring independently or by copying other individuals.

Second, we find substantial variation in the types of copying behavior that different individuals display. Some individuals appear to focus their copying behavior

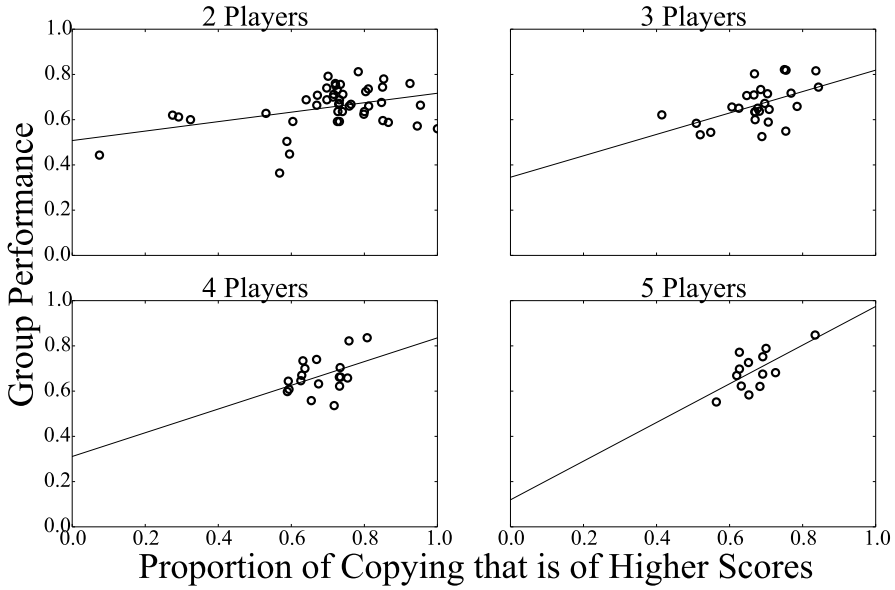

Figure 3-5: Average group performance as a function of the fraction of copying in the group that consists of "intelligent copying"—copying of an individual with a higher score. Lines are individually fitted regression lines.

on other players who tend to have higher scores, whereas other individuals appear to be less discriminating in their copying behavior. Moreover, as shown in Figure 3-5, groups that contain individuals who focus their copying behavior on higher scoring individuals achieve significantly higher performance in our task (slope: 0.2639, 95% CI: $[0.145, 0.383]$). This result, though subject to the confounding of correlation and causation, could be explained by theory of mind assisting in individual and group performance. A player who is able to accurately infer whether another player is receiving a high score may be able to achieve higher performance on our task by leveraging these inferences to more effectively copy others.

3.4 Behavioral Model

The trends we observe suggest a potential set of behavioral mechanisms that effective human groups may use in our task. We propose that each player in an effective group chooses a state based on the following rules:

1. If the player is in a good area, the player will remain in that area exploiting.

2. If the player in not in a good area and the player perceives another person as possibly having a higher score, the player may choose to copy that person.

3. Otherwise the player will explore independently.

According to this model, players in bad locations improve their scores by copying exploiting individuals instead of wasting time by copying low scoring players or wasting time by exploring many poor quality areas. The model also has interesting emergent collective properties. When any individual finds a good area, that player will attract the other players to that location by exploiting. Then, when all the players are together in a group exploiting a particular area, one of the players will start to lose bonus as the score field shifts. This player will then either move closer to the others who are still exploiting or will shift to an exploring state. If that player starts exploring but doesn't find any good locations, the player will return to the group if the group is still exploiting. If that player does find a new good area, though, the player will start exploiting that area. The rest of the group will then follow after the highest scoring region shifts to where the exploiting player is. This mechanism creates a kind of gradual crawling that effectively tracks the moving score field. Thus, by using this mechanism players are improving both their own performances directly and also that of the entire group by participating in this process of emergent collective sensing. An example of this process occurring in participant gameplay is shown in Figure 3-6.

3.5 Discussion

In our experiment, we observed that humans were able to achieve increases in performance at much smaller group sizes than fish. Fish exhibited mild improvements in group performance at groups of 16 and more substantial improvements at groups of 64 and 128. However, we see significant improvements in human performance at just five players. This difference may be at least partially explained by the differences in the mechanism that humans appear to use in this task as compared to fish.

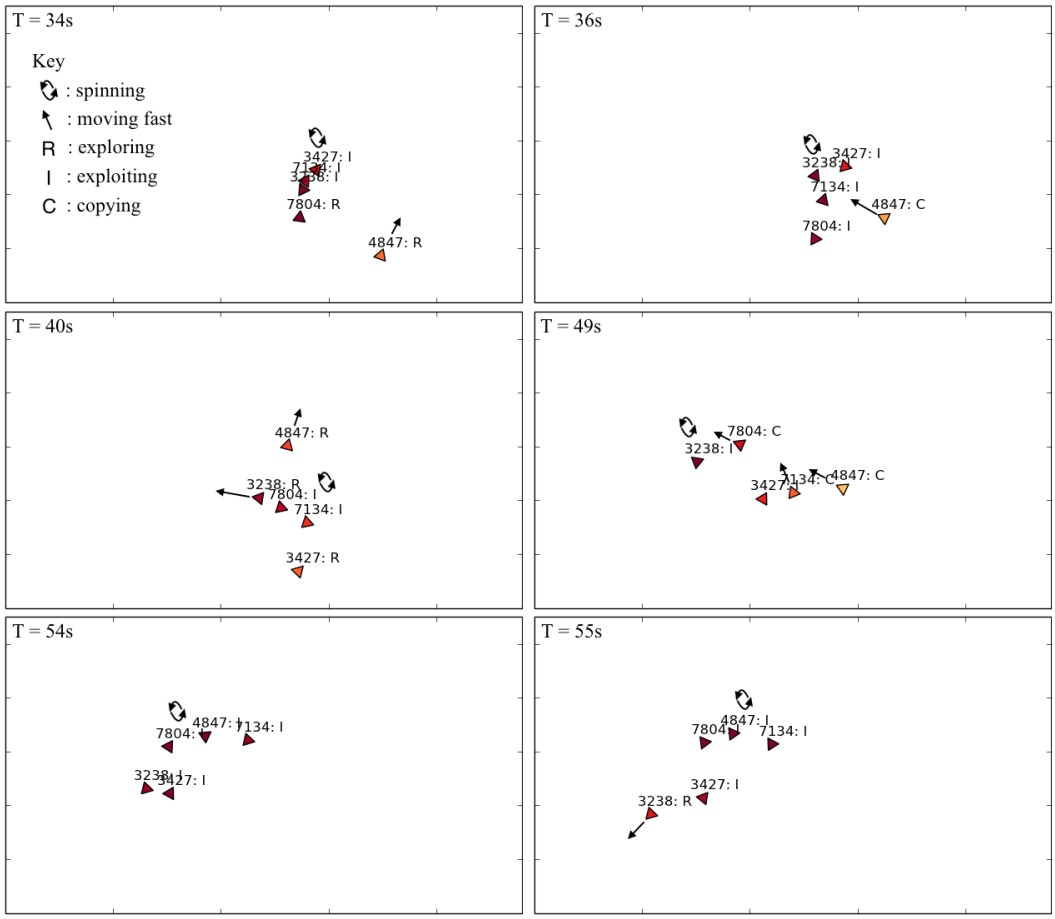

Figure 3-6: Reconstructions of actual gameplay in a five-person group illustrating both failed exploration leading to intelligent copying and successful exploration leading to collective movement. Colors indicate the individuals' scores, with red being higher and orange/yellow being lower. The player labels indicate both player IDs and also the player states our feature extraction procedure inferred. Other annotations are provided to give a sense for the game dynamics. At 34 seconds, in the first panel, most of the group has converged on exploiting a particular area while one individual is exploring independently. To the right, at 36 seconds, the exploring individual appears to have failed to find a good location and ceases exploring by copying the group. At 40 seconds, the final panel in the first row, the score field has shifted and some of the group begins exploring while others continue to exploit. By 49 seconds, the first panel in the second row, one of the exploring individuals found a good location, and other players have begun to move towards that individual. At 54 seconds, the entire group is exploiting the new area. In the final panel, at 55 seconds, the background has shifted enough again that one of the individuals begins to explore.

Interestingly, the mechanism we identify in humans is similar to that of fish in some ways, but it is also distinct important ways. Similar to the behavior of humans in choosing appropriate states based on current score, fish modulated their speeds based on the level of darkness that they were experiencing. Fish moved slower in their preferred darker areas and faster in lighter areas. Similar to the copying behavior we observe, fish had a tendency for turning towards other fish. However, Berdahl et al.'s model of the behavior of their fish did not require any reference to the kind of discerning social awareness that we see in humans. Whereas fish appear to equally weight information from all nearby conspecifics, effective humans appear to modulate their copying behavior based on the inferred scores of other players. The strategic use of independent exploration (a form of asocial learning) was also key to the mechanism enabling human success. These key differences support recent work in social learning [125, 82], which find an impressive flexibility in the strategic deployment of imitation in humans. Of course, it is difficult to compare human performance directly to that of fish given the differences between the perceptual and motor abilities of fish in an actual fish tank and the abilities of the participants in our simulated environment. Nevertheless, our comparison hints at a superior capacity for distributed cognition in humans, possibly enabled by our ability for theory of mind.

Perhaps an even more interesting difference that emerged between humans and fish has to do with the time scale over which the collective intelligence mechanism evolved. For fish, the ability to gain from group performance in these collective sensing tasks is likely based on innate behaviors, selected over many generations of fish facing exactly this problem over their whole lifespans. In contrast, some of our humans groups, facing this particular problem for the first time, appear to have discovered reasonable collective sensing strategies in just a matter of minutes.

Beyond the recent literature on collective intelligence in nonhuman animal groups, there has been a long line of work studying the factors that predict the performance of human groups in various scenarios [62]. Our findings are consistent with previous work suggesting that having a larger group is beneficial in complex, uncertain environments [107]. Unlike much of this previous work, however, we focus here on the possibility in

larger groups of new emergent group abilities and behaviors, and on the mechanisms leading to these emergent properties.

This work therefore may shed light on one of the pressing puzzles of human collective intelligence and human distributed cognition. What are the specific mechanisms by which humans establish effective coordinated distributed information processing agents that can accomplish more than any individual alone, and how do our abilities play a role in these mechanisms? The perspective of group behavior as distributed processing [55] suggests the importance of communication for collective intelligence because of the importance of communication in distributed systems. Moreover, theory of mind—an enabler of implicit communication—has been shown to be predictive of collective intelligence [126, 31]. While this work does not have a powerful enough experimental design to be definitive, this work at least further suggests that one of the roles that theory of mind plays in the emergence of collective intelligence is facilitating implicit communication that allows for coordination on good collective actions. Moreover, this work also suggests that the benefit of a group's coordinating on good actions could be more than simply the benefit to each individual independently. By combining a natural human tendency for independent exploration with a discerning social awareness, humans appear to be able to fluctuate between exploiting known good actions, independently exploring new options, and intelligently copying the promising choices of other individuals. A simultaneous combination of these activities by a cohesive group appears to lead to a collective memory of recently good actions from individuals who continue to exploit, and a collective movement towards actions that promise to be good in the near future driven by independently exploring individuals. The reactive distributed sensing ability that appears to emerge from this process may confer a unique benefit to working together in tightly knit groups.

Chapter 4

Information Aggregation in Shared Belief Formation

So far we have seen that in a stylized example, collective agency can emerge from a simple social imitation mechanism. In this chapter, I investigate a richer example [69] from a real-world data set. From deciding how to invest to deciding what career to choose, people are faced with many complex decisions among large sets of options. These decisions are often further complicated by the fact that new information is constantly arising about the large number of options available. How do boundedly rational people synthesize this overwhelming amount of information to make well-informed decisions? Previous research has shown that people have a strong propensity to imitate decisions that others make [7, 25, 89]. However, the aggregation properties of these imitative heuristics are sensitive and poorly understood—too much or too little imitation can lead to failures in information aggregation [94], and it is not known if or how people balance this trade-off. I use a unique large-scale behavioral dataset from an online social network of 50,000 amateur investors to provide evidence that people engage in an imitative heuristic equivalent to directly incorporating decision popularity as a naïve Bayesian prior distribution in shared decision-making settings. This heuristic, while boundedly rational from an individual perspective, collectively balances exploration and exploitation to achieve highly effective, and evolutionarily adaptive, information aggregation: I show that the heuristic implicitly implements

a near-optimal Bayesian decision algorithm called Thompson sampling [114]. By revealing this relationship between heuristic imitative behavior and collective Bayesian belief formation, I illuminate how the emergence of rational collective learning can be precisely characterized as a kind of distributed Bayesian inference. At the same time, the heuristic behavior I observe can be fragile, suggesting that the structure of the sociotechnical systems in which decision-making occurs can promote or undermine rational collective belief formation.

4.1 Introduction

There are thousands of companies listed on the New York Stock Exchange and the NASDAQ alone. The number of options a person has for what career to choose or what path to take in life are uncountable. Even in decisions as mundane as where to buy a cup of coffee or where to go out to eat dinner, a city dweller is faced with a dizzying array of options—Boston's North End neighborhood has over 50 Italian restaurants; downtown Manhattan has hundreds of bars. Furthermore, the information available about the options in each of these cases changes over time. News and financial statements about companies are released daily. Certain careers or life paths appear more or less tenable in different world conditions. New reviews are written about cafés, restaurants, and bars every day. We are faced with innumerable complex decisions over the course of life.

Fortunately, we do not face these decisions alone. Entire communities of people are faced with the same sets of options, and can communicate information about different options available. Decision-making with the benefit of the accumulated knowledge of a community can result in far superior decisions compared to what people could achieve alone [99, 50, 15, 80]. Yet, groups still face two coupled challenges in accumulating knowledge to make good decisions: aggregating information, and addressing an informational public goods problem known as the exploration versus exploitation dilemma [78, 117, 51]. Ideally, we could all pool the experiences we have had to determine the best possible rational beliefs about the qualities of the

options available. Obviously a naïve implementation of this ideal—directly sharing all personal preferences and experiences—is not achievable due to limitations in human memory, communication bandwidth, and self-awareness. The second fundamental challenge occurs even if we were able to do ideal information aggregation. If everyone focuses on the best-looking options at a given time according to all the available information available, we learn little about less-explored potentially better options.

The science on how human groups address these challenges is scattered. A huge body of empirical work shows that people imitate each others' decisions in a wide range of scenarios [102, 46, 89, 118, 24], and countless mathematical models of how social signals affect decision-making have been proposed. A subset of these models explicitly deal with how people in groups process information available [14, 21]. Some of these information processing models, as well as others that do not explicitly represent information processing, examine whether groups eventually converge on good or bad decisions [13, 40]. However, this asymptotic reasoning leaves a gap between the mathematical analysis and our day-to-day experiences. Showing that a decision-making mechanism eventually leads to a good decision says nothing about how long the process takes to get there. Eventually identifying a good decision is quite different from making the best decision you can at each step of the way. Relatedly, these prior models fail to address how well groups are able solve the exploration-exploitation dilemma. Other models have been proposed to reason about exploration-exploitation [72, 80], but sidestep the issue of information aggregation by focusing on discovery of options with known quality rather than requiring repetition to identify the true quality of options.

Here I propose a model that accurately captures how people perform information aggregation and also balance exploration with exploitation in millions of decisions performed within an online social financial trading platform. This model reveals that ideal groups can aggregate information efficiently, while collectively balancing exploration and exploitation, by using a simple heuristic probabilistic decision-making strategy that approximately implements a near-optimal Bayesian learning and decision-making algorithm in the aggregate. In other words, I show that groups can

achieve highly effective decision-making as a community without requiring individuals to do any sophisticated calculations. Individuals use a simple rule that acts like a component of a distributed algorithm. The computational properties of this distributed algorithm precisely characterize the information processing abilities of the group. Even though individuals are not fully Bayesian themselves, the community as a whole can implement a Bayesian decision algorithm. The advantages of this decision-making strategy lead it to be favored in evolutionary simulations. These results shed light on the mechanisms, properties, and evolution of human collective intelligence.

4.2 Environment Model

We consider a widely used abstract representation of large decision-making problems in order to formulate our model. We represent people as making decisions among a set of options that probabilistically generate rewards over time. These rewards correspond to the benefits that a person who takes an option at a particular time would obtain—for example, the performance of a stock on a particular day. For simplicity in developing our model, we assume that a fixed set of M options is available over time, that rewards are perceived just as either good or bad (represented as 1 or -1), and that rewards are generated independently at random. Mathematically, each option j is associated with a probability θ_j of generating a positive reward, x_{jt}, at each time t: $P(x_{jt} = 1) = \theta_j$. In the idealized case, each option produces a reward at each discrete time step—for example, each day. The historical rewards of an option up to a particular time t forms a set of information about that option, $\boldsymbol{x}_{j,\leq t} = \{x_{j1}, x_{j2}, \ldots, x_{jt}\}$. We denote the set of information about all options at time t as $\boldsymbol{X}_{\leq t} = \{\boldsymbol{x}_{1,\leq t}, \boldsymbol{x}_{2,\leq t}, \ldots, \boldsymbol{x}_{M,\leq t}\}$. In engineering, an environment satisfying this set of assumptions is commonly referred to as a stochastic multiarmed bandit. The "single-agent" case, in which only a single decision-maker is present in the environment, is a well-studied problem in computer science and discrete optimization, with important applications ranging from medical trials [119] to online advertising [59]. The "multi-

agent" case is a growing area of interest in these communities.

4.3 Problem Formulation

Effective decision-making in this environment requires maintaining representations of information available. Bayesian posterior distributions are information-theoretically optimal representations [91]. One way to formulate the information aggregation problem is therefore as computing the Bayesian posterior distributions over the reward probabilities of each option, $P(\theta_j \mid \boldsymbol{x}_{j,\leq t})$. One main challenge comes when different people observe different pieces of information. If each person observes a subset of $\boldsymbol{X}_{\leq t}$, then people must communicate to combine their information in order to base their decisions on all the information that the group has available. The problem of information aggregation is therefore to establish a communication procedure that enables each decision-maker to compute all the full posterior distributions, $P(\theta_j \mid \boldsymbol{x}_{j,\leq t})$, from the information locally available to each decision-maker.

There is also another challenge in this sort of environment. Just being able to aggregate and represent information is not enough to achieve effective decision-making. There is also a problem of collecting useful information. Given access to a representaiton of the information a group has available, it is straightforward for a person to make the best-looking decision given that information. However, this "best-looking" decision may not be the best decision possible. In order for individuals to make good decisions compared to the best possible decisions they could be making, the group must continually gather information to learn about options of uncertain quality. At the same time, people should not suffer an undue opportunity cost by exploring uncertain options too much. This challenge is referred to as the exploration-exploitation dilemma [111].

A great deal is known about computational solutions to the engineering problem of the single-agent exploration-exploitation dilemma in stochastic multi-armed bandit problems. Upper bounds on achievable performance have been discovered, and provably optimal algorithms exist that meet these upper bounds. One popu-

lar algorithm for solving multi-armed bandits that works well in practice [22] and was recently proven to be at least near-optimal [1, 58] is called Thompson sampling. The Thompson sampling algorithm maintains representations of the Bayesian posterior distributions over the quality of each option available. To make a decision on a particular time step, the algorithm probabilistically samples an option with probability equal to the probability that option is the best option available, $P(j \text{ is best} \,|\, \boldsymbol{X}_{\leq t}) = P(\theta_j > \theta_k \ \forall k \,|\, \boldsymbol{X}_{\leq t})$. Of course, the single-agent case does not involve aggregating information across individuals, and hence does not solve both challenges associated with decision-making in the multi-agent case.

The extension of Thompson sampling to the context of a group of decision-makers, which couples the exploration-exploitation dilemma with the information aggregation problem, remains an active area of research [50]. One potential extension of single-agent Thompson sampling is for individuals to somehow aggregate information so they can implement Thompson sampling on all of the information the group has obtained, rather than just the information that the individual has collected personally. If people can achieve this "group Thompson sampling" via some information aggregation procedure, they will essentially be making decisions at least as effectively as an individual could with the benefit of all the accumulated knowledge of the group.

This possbile solution to the dual challenges of the collective exploration-exploitation dilemma and the information aggregation problem therefore involves accumulating knowledge into Bayesian beliefs, and using these beliefs in order to implement a group Thompson sampling procedure. While simple to state, achieving these aims appears difficult for boundedly rational individuals, who might have too limited memory to store all the information a group has received, too limited processing power to use it all, and potentially too limited communication bandwidth to even receive it all. A large group faced with a large number of options will accumulate a huge amount of information, and Bayesian posterior distributions can be computationally costly to compute. The question we proceed to investigate is whether it is possible for a group of boundedly rational individuals to achieve this form of rational collective behavior.

4.4 Behavioral Model

We introduce a behavioral model of human decision-making in the context of a group of decision-makers faced with a shared set of options, and we compare this behavior to the normative standard of group Thompson sampling. Our approach to behavioral modeling is to start with existing descriptive accounts of human decision-making, refine these accounts with the benefit of modern computational cognitive theory, and reinterpret them in light of our normative standard. Strong evidence for specific behavioral models of group decision-making is still relatively sparse in the existing literature, but there is literature based on the evidence that we have that identifies plausible decision mechanisms people might use for decision-making in group contexts.

As a starting point, we observe that two-stage decision-making models are common [54, 93, 70, 103, 96]. These models suppose that people filter the options they might want to consider according to some initial criteria, and then choose which decision to make among this smaller set of options according to another set of criteria. The mechanism we propose is a member of this class of models. As in some of these related models [70, 103, 96], we suppose that a person first selects an option to consider according to the current popularity of that option, and then decides whether to commit to that option according to a recent signal of the option's quality. The model we propose refines the existing accounts that have this structure with a more cognitively grounded procedure in the second step. Our key technical refinement allows us for the first time to incorporate a specific account of how people use information signals in their decision-making, which was lacking in these prior two-stage models. We are thereby able to understand the information processing features of this type of two-stage decision-making heuristic, both in terms of information aggregation and collective learning.

Specifically, the model we propose, which we call "social sampling", consists of the following two-stage procedure. A person who wants to make a new decision first selects an option to consider at random with probability proportional to the current popularity of that option, p_{jt}. A natural way for this first step to implemented is

to choose a person who has already made the decision uniformly at random, and to ask about what option that person selected. The step could also be implemented via a stochastic response to a list of options ranked according to popularity. This step reduces the cognitive burden of evaluating many options by allowing the decision-maker to consider only a small set of options, rather than all the options available.

The second step is for the person to decide whether to commit to the initially selected option. Here we propose that the person performs an abbreviated Bayesian computation to assess the quality of the option being considered. To formulate this second step, we suppose that each person has an idea about what sort of rewards constitute a good option in a particular decision at hand. That is, people have certain expectations about what type of payoffs good options will produce, and how often. The way this assumption manifests in our behavioral model is that people suppose that a "good" option is one that has a certain option-invariant probability of producing positive rewards, $P_i(x_{jt} = 1 \,|\, j \text{ is good}) = \eta_i$. This probability might vary from person to person, as indicated by the index i, but is shared for that person by all options that are considered good.

In this second step, each person observes a recent quality signal x_{jt} associated with the option that person has sampled to consider during this step. Each person then assesses how likely that signal is given that the option is good, and commits to the option if the signal is likely assuming the option is good. This step is analogous to probability matching on the posterior that the option being considered is good, $P_i(j \text{ is good} \,|\, x_{jt})$, but ignores the normalizing constant $P_i(x_{jt})$. This step is motivated by recent results in the cognitive science literature arguing that people resort to approximate Bayesian computations in many decision-making scenarios [120, 35]. Normalizing constants can be difficult to compute. Likelihood-based sampling strategies circumvent this computational burden, while also helping account for stochasticity observed in human behavioral data.

In summary, the social sampling model supposes that people select options to consider by consulting others' decisions, and then commit to options being considered by privately evaluating whether the options seem good according to recent information

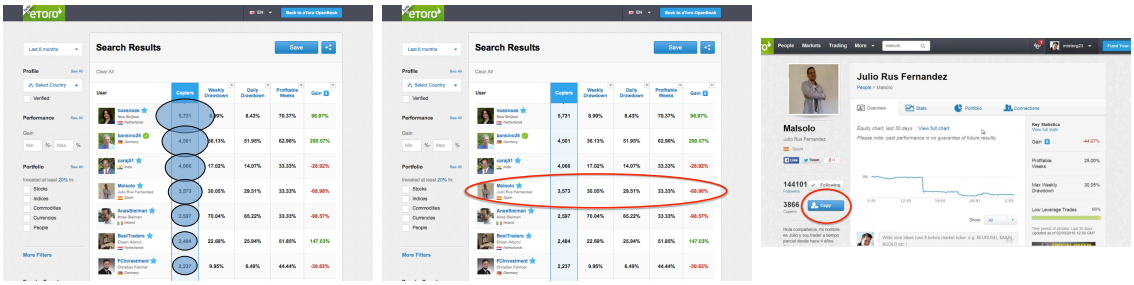

Figure 4-1: **Social sampling can easily be implemented on eToro.** This figure displays images of the eToro interface illustrating how social sampling could be implemented on eToro. The search interface is the most prominent mechanism for finding users on eToro. Social sampling could be implemented by users sorting by popularity, probabilistically choosing a trader to consider according to that list, then assessing the recent performance of that user more careful before deciding to follow the user. The fact that the eToro interface affords users the ability to implement social sampling does not diminish the importance or generalizability of the mechanism. Similar mechanisms have been hypothesized across a range of species and contexts.

available. Together, these steps yield the following probability of a person committing to a particular option j at time t:

$$\frac{p_{jt}^{\gamma} P_i(x_{jt} \mid j \text{ is good})}{\sum_k p_{kt}^{\gamma} P_i(x_{kt} \mid k \text{ is good})} = \frac{p_{jt}^{\gamma} \eta_i^{x_{jt}} (1-\eta_i)^{1-x_{jt}}}{\sum_k p_{kt}^{\gamma} \eta_i^{x_{kt}} (1-\eta_i)^{1-x_{kt}}}.$$

The γ parameter determines how much weight people give to popularity. When $\gamma = 0$, people ignore popularity entirely. When $\gamma \to \infty$, people always decide on the most popular option.

4.5 Evidence

To test this model, we use a large-scale dataset from users of an online social network of amateur investors called eToro.[92] eToro's platforms allows users to make trades on their own, predominantly in foreign exchange markets, or to choose other users on the site to follow. When one user chooses to follow another, the follower allocates a fixed amount of funds to automatically mirroring the trades that the followed user makes. eToro then proportionally executes all of the trades of the followed user on the follower's behalf. A substantial fraction of people on eToro use this social

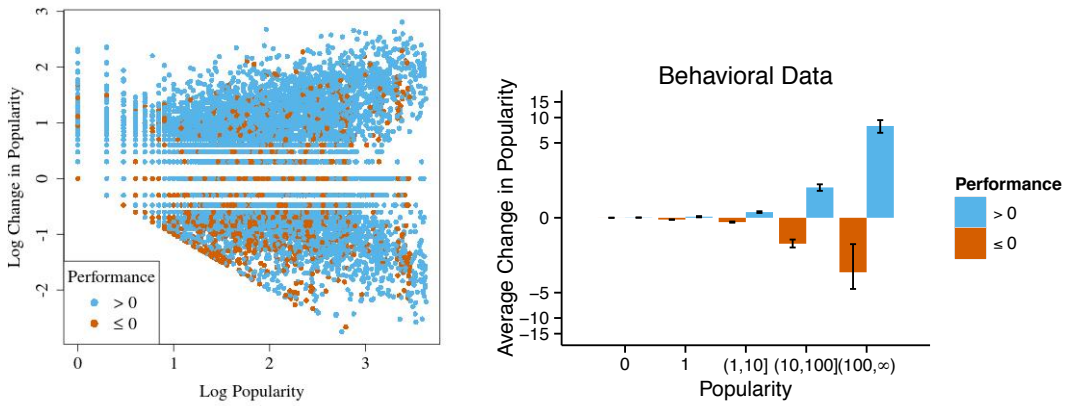

Figure 4-2: **Daily changes in popularity on eToro tend to be positive for those traders who are performing well and negative for those traders who are performing poorly, and the magnitude of those changes are greater as popularity increases.** (Left) A scatter plot illustrating the observed relationship between daily change in popularity on eToro with past popularity and recent performance. There is one data point shown for each trader on each day. Points are colored by whether recent performance is positive or negative. (Right) A bar plot visualizing the same data to highlight the trends we model.

investing functionality. At the time our data was collected, the site had over 50,000 users. People are hence faced with a large number of options of who to follow. In order to assist in users' decisions about who to follow, eToro provides information about the trades and trading performance of each user via a search interface and public profiles. These statistics vary over time and must be carefully integrated to form a complete picture of each trader's performance. Users on eToro are therefore faced with a difficult decision problem of choosing who to follow among a large set of options, given unreliable, multidimensional, temporally varying performance signals. eToro also presents the current popularity of each user, i.e., the number of people currently following that user. A plausible way the social sampling mechanism could be used on eToro via these social and non-social signals is illustrated in Figure 4-1.

To map the eToro context onto our abstract environment model, we treat each user on eToro as an "option" of someone who could be followed. We analyze the data at a daily temporal granularity, and we model the new follow decisions each user makes on each day. We summarize trading performance with expected return on investment

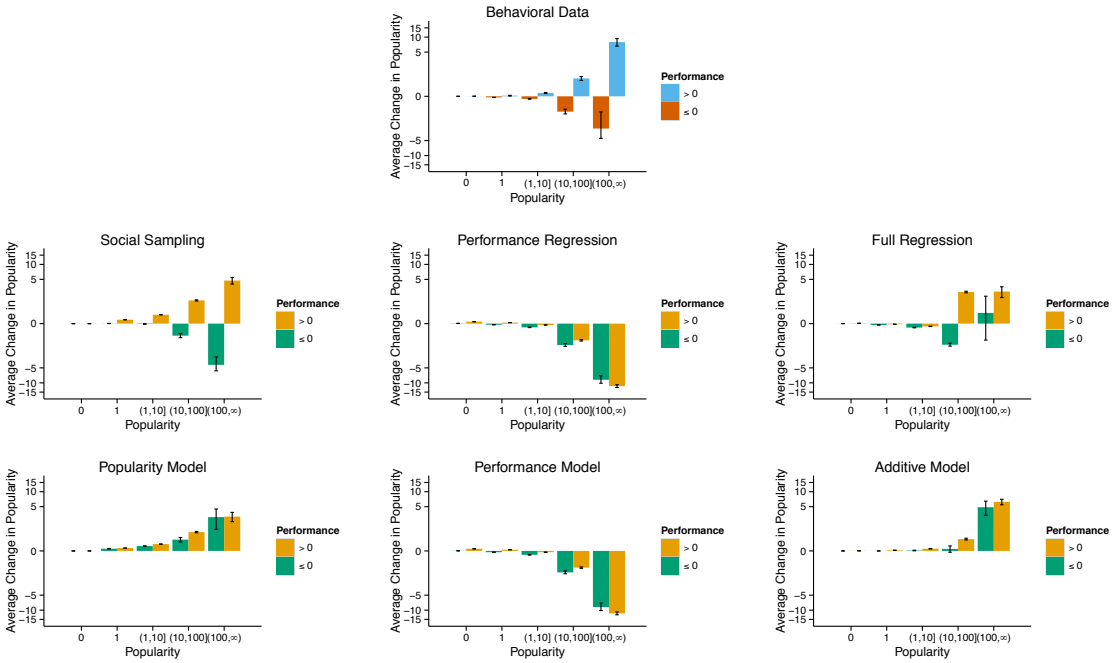

Figure 4-3: **Social sampling replicates the observed patterns in eToro's popularity dynamics better than several plausible alternative models.** We retrospectively predicted how new follow decisions would be distributed according to social sampling, as well as according to several alternative models. Models that rely more on performance information tend to underestimate the impact of popularity, while models that rely more on popularity tend to underestimate the impact of performance.

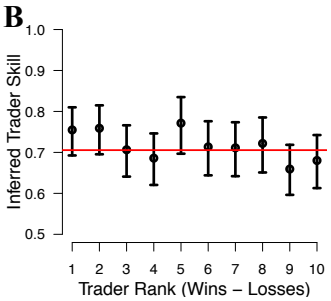

Figure 4-4: **Our model fit provides a consistent estimate of the performance of the best traders on eToro.** The fitted value of η is shown along with direct estimates of the probability each of the best traders on eToro has of making winning trades.

(ROI) from closed trades over a 30-day rolling window, which is similar to the "gain" performance metric presented to the site's users. We consider the reward signal associated with following a person on a particular day to be positive if that person's performance is greater than zero that particular day, and negative if performance is less than zero.

We first observe that as predicted by the social sampling model, people appear to rely heavily on popularity in deciding who to follow (Figure 4-2). To observe this pattern, we look at changes in the popularity of each user from one day to the next. Users who perform well tend to gain followers, and users who perform poorly tend to lose followers. At the same time, the magnitude of these changes becomes larger as the popularity of the user increases. People with few followers are unlikely to gain many followers, even when they perform well. People with higher popularity gain more new followers when they perform well, but lose more followers when they perform poorly.

For a more rigorous test of the social sampling model, we fit the parameters of the model and compare the quantitative predictions of the model to those of several plausible alternative models. When fitting the social sampling model, we assume $\eta_i = \eta$ for a fixed η since this collapsed model yields the same aggregate behavior in the idealized case. We examine five alternative models: A "performance regression" attempts to predict follow decisions based on a logistic regression model using the ROI performance metric as a dependent variable. A "full regression" attempts to predict follow decisions based on a logistic regression model using the ROI performance metric, popularity, and a multiplicative interaction term between these two signals as dependent variables. The "popularity model" corresponds to the social sampling model with $\eta = 0.5$, which neglects performance signals completely. The "performance model" corresponds to the social sampling model with $\gamma = 0$, which neglects popularity information. The "additive model" corresponds to an additive mixture of the popularity and the performance model (as compared to the social sampling model, which is a multiplicative combintion of these two "lesion" models). After fitting these models according to the new follow decisions that users on eToro make, we

find that only the social sampling model is able to reproduce the empirical patterns in the dynamics of popularity from our behavioral data.

We also examine the internal consistency of the fitted social sampling model. The η parameter in the social sampling model represents the probability that a trader will generate positive performance signal on a particular day. Since we fitted this parameter directly to the follow decisions that were made, rather than to the empirical probabilities of traders achieving positive performance signals, we can examine if the fitted η value corresponds to these empirical probabilities. We find that the fitted η value does indeed correspond closely to the probability of the best traders achieving positive performance (Figure 4-4).

Anecdotal reports from users of eToro also corroborate the social sampling model. One website states: "Here are some tips and things we look at when selecting the Professional Investors/Traders we copy: Most people will still want to start with looking at the 'most copied traders'... Popularity is obviously a decent 'starting' point for finding traders to analyse further... putting in some time and effort to analyse the additional statistics is likely to lead to better long term results."[1] Another user explains: "[Ranking by gain] is what eToro's [standard] ranking system is showing. And but this is not so much the way to really choose who's a good trader because you just don't know how long these traders have been trading until you go into details. Another good idea is to do this: What I'll do is, you just go to 'Copiers', you just select the 'Copiers' tab and show the ones who have the most copiers. Now, this isn't a full-on good guide either, just choosing the amount of copiers. Cuz as you see, there are other traders in this line who have been making more. Like an example, over 300%, over 300%, and they're down the line just cuz they've got less copiers."[2] These users are both describing how to use popularity as a kind of first-pass filter, before looking at performance information. Social sampling is clearly an intuitive and appealing decision-making heuristic.

[1] "eToro Tips: Find Best Gurus" from SocialTradingGuru.com (http://socialtradingguru.com/tips/etoro-tips/select-etoro-gurus).

[2] "How to Choose Good Traders or Gurus to Copy in Etoro? - Video Guide" YouTube (https://www.youtube.com/watch?v=di7Sw587has, 3:02)

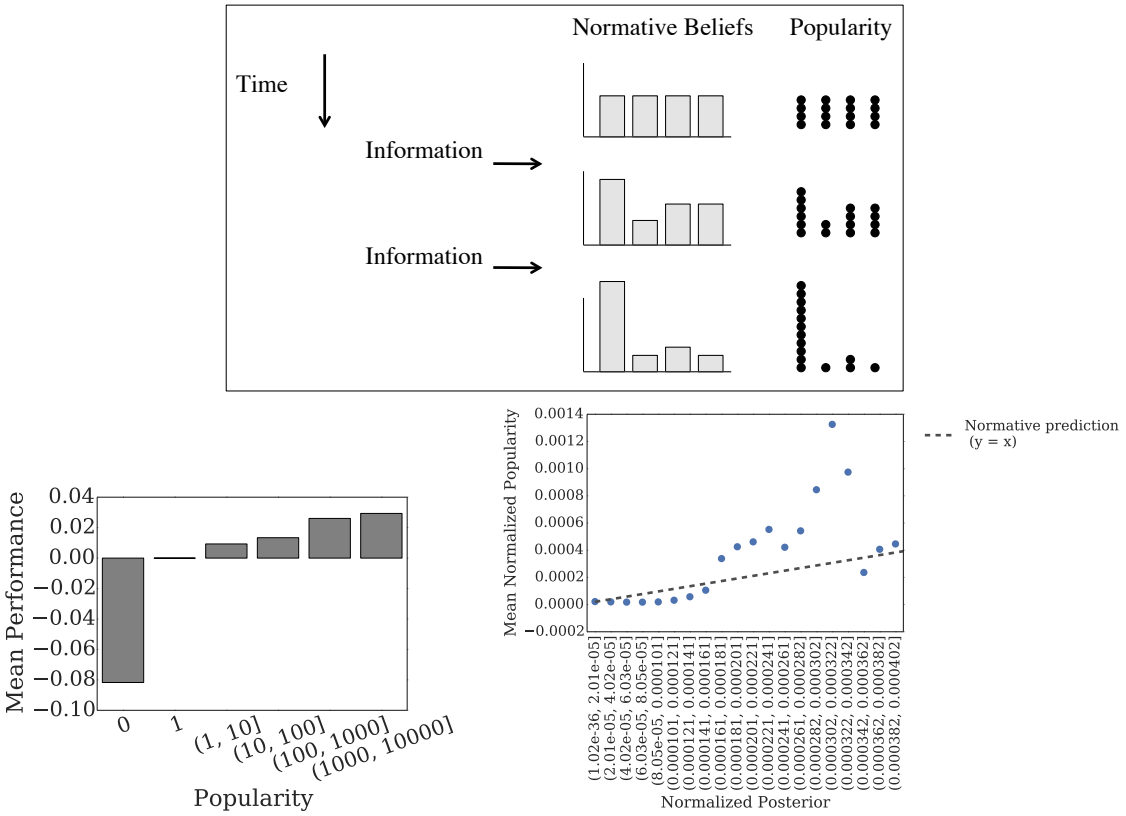

Figure 4-5: **The social sampling model predicts that popularity should track a normative posterior distribution, and this prediction is borne out in our data.** (Top) An illustration of the predicted relationship between popularity and the normative beliefs about which traders are good. (Left) A plot showing the general relationship between popularity and performance. (Right) A plot showing the match between the normalized posterior and popularity on eToro. A linear regression between normalized popularity and the normative posterior has a fitted value of about 1.5 ($p < 0.001$), which is quite close to the predicted value of 1.0.

4.6 Model Analysis

Remarkably, despite it's simplicity and heuristic appearance, the social sampling model achieves both excellent information aggregation and a highly efficient balance between exploration and exploitation. In order to analyze the social sampling model, we consider a simplified model of the world. In this simplified model, known as a "hide-and-seek" problem [104], there is a single best option that has a probability η^* of producing positive rewards, while all other options produce positive and negative rewards uniformly at random. When the number of options is large or $\eta^* = 0.5 + \epsilon$ for

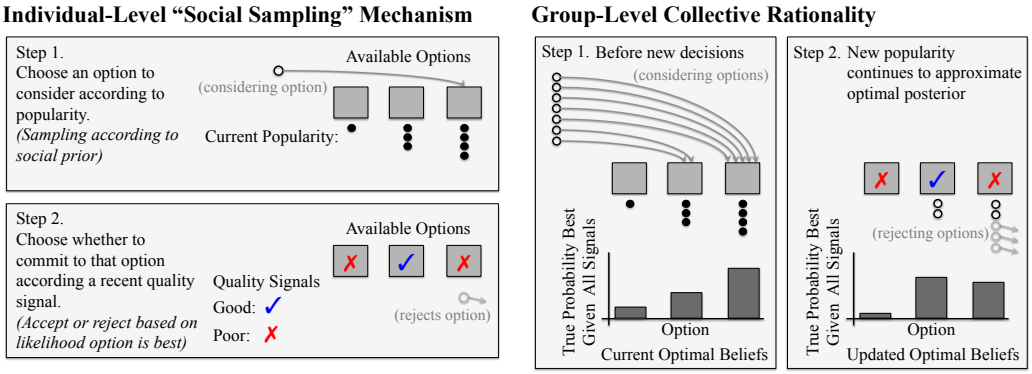

Figure 4-6: A schematic illustration of the "social sampling" mechanism we propose as a model of human social decision-making, and an illustration of how this mechanism yields collectively rational belief formation at the group level. Individuals treat current popularity as a prior distribution, and sample according to this prior in order to choose an option to consider taking. An individual then commits to a considered option with probability proportional to the likelihood that the option is best given a recent objective signal of quality. If current popularity approximates the current optimal posterior distribution that each option is best given all previous quality signals, then when a large group of decision-makers continues to use the social sampling mechanism, popularity will continue to approximate the optimal posterior.

small $\epsilon > 0$, this hide-and-seek setting can be thought of as a pessimistic assumption about the identifiability of the best option in the environment. In other words, this setting can be interpreted as one in which good options are rare or difficult to identify. A companion piece of work shows that our results also generalize to other contexts [19].

In this case, the Bayesian posterior required for Thompson sampling simplifies to $P(j \text{ is best} \,|\, \boldsymbol{X}_{\leq t}) = \frac{(\eta^*)^{x_{jt}}(1-\eta^*)^{1-x_{jt}} P(j \text{ is best} \,|\, \boldsymbol{X}_{<t})}{\sum_k (\eta^*)^{x_{kt}}(1-\eta^*)^{1-x_{kt}} P(k \text{ is best} \,|\, \boldsymbol{X}_{<t})}$. This probability bears a striking resemblance to the probability of choosing option j under social sampling. In fact, when $\gamma = 1$ in an infinite population of decision-makers who implement social sampling using $\eta_i = \eta^*$, the following invariant will be maintained: $p_{jt} = P(j \text{ is best} \,|\, \boldsymbol{X}_{<t})$. Popularity can therefore be precisely understood as compactly summarizing the past information about the options available to decision-makers under a pessimistic assumption about the environment. In other words, popularity under social sampling with $\gamma = 1$ has an exact correspondence to a Bayesian posterior distribution that each option is best, and social sampling leads decision-

makers to sample exactly according to the group Thompson sampling probability $P(j \text{ is best} \,|\, \boldsymbol{X}_{\leq t})$, with the benefit of all of the information the group has accumulated.

4.7 Theoretical Predictions

The theoretical results of our model analysis generate sharp testable predictions. First, we can compute the posteriors $P(j \text{ is best} \,|\, \boldsymbol{X}_{\leq t})$ empirically and test whether popularity tracks these distributions. To compute these posteriors in the complex environment of the eToro data, we relax the assumption that there is only a single best option, so that $P(j \text{ is best} \,|\, \boldsymbol{X}_{\leq t}) = P(j \text{ is good} \,|\, \boldsymbol{X}_{\leq t}) = P(j \text{ is good} \,|\, \boldsymbol{x}_{j,\leq t})$. Without this relaxation, computing the posterior is complicated by missing data on days without observations. We find that not only is popularity generally related to performance, popularity on average matches these normative posterior values remarkably well.

The relationship between social sampling and group Thompson sampling also indicates that social sampling should be able to achieve a highly effective balance between exploration and exploitation, leading to fast discovery of good options and high overall performance. However, since the relationship requires $\gamma = 1$, these features may require that popularity is weighted linearly as opposed to superlinearly or sublinearly. We do find in idealized simulations that social sampling achieves excellent performance when $\gamma = 1$ and that this performance declines quickly when $\gamma < 1$ or $\gamma > 1$.

4.8 Implications for Design

Our results have implications for the design of sociotechnical systems. On eToro, we would expect traders to make the most money from following others if they are using social sampling with $\gamma = 1$. Remarkably, we find that users on eToro sometimes achieve this ideal when we fit the γ parameter over time, to the subset of follow

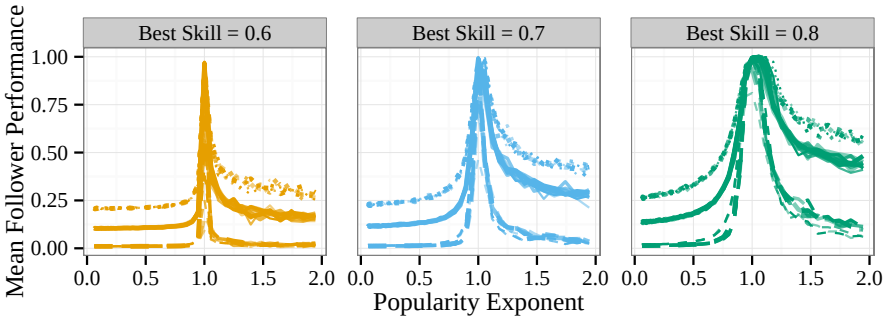

Figure 4-7: **Social sampling with $\gamma = 1$ is far more effective than sampling sampling with sublinear or superlinear use of popularity, suggesting that the Bayesian interpretation of social sampling is critical to its performance.** An idealized simulation experiments. Each line represents a different combination of simulation parameters, with the four graphical parameters of the plotted lines (line type, width, color, and transparency) each representing a simulation parameter.

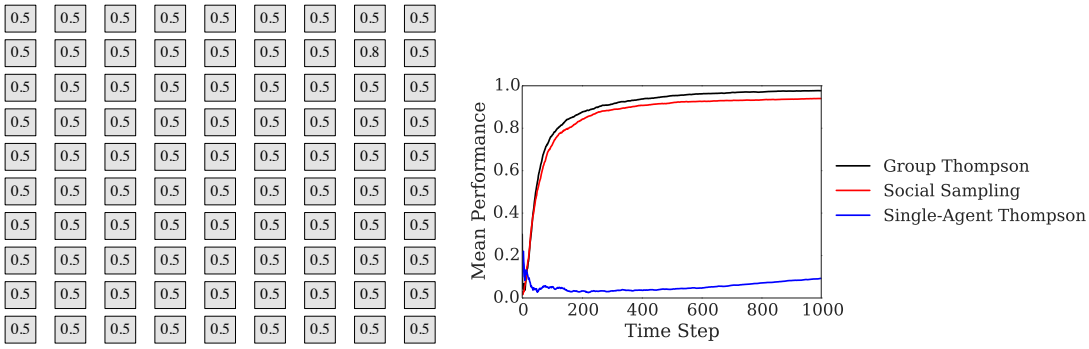

Figure 4-8: **Social sampling far outperforms how quickly an individual could learn alone, and nearly achieves the performance of group Thompson sampling.** (Left) An schematic illustration of the difficulty of finding the best option among a large set of options. One box has a number greater than 0.5. (Right) Results from an idealized simulations using 100 options and 10,000 social sampling agents to illustrate differences in convergence time.

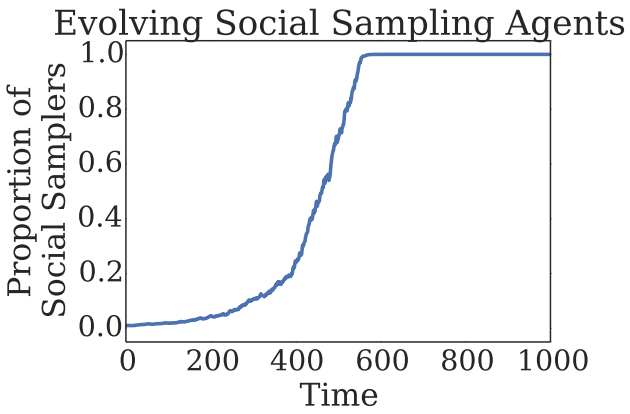

Figure 4-9: **Social sampling is favored over individual decision-making in an evolutionary simulation.** Starting from a population of 99% non-social agents and 1% social sampling agents, social sampling agents gradually grow to overcome the population in an evolutionary simulation.

decisions made on each day. However, on many days—and if we fit γ on all the data together—we find that γ is usually fitted to be less than one. Our retrospective simulations indicate that the overall fitted value of about 0.73 is highly suboptimal for the population as a whole. However, this value does not do much worse than the optimal value of $\gamma = 1$ when considering the actual observed popularity values on eToro, rather than ideal values in a population of social samplers. Therefore, people are roughly individually best-responding to the social signals in their environment, but there is a systemic problem, perhaps caused by variability in the set of near-best response strategies. These findings indicate that eToro could better serve its users by altering its interface to encourage users to utilize popularity linearly, rather than sublinearly. More generally, these results show how the efficiency of collective learning in an environment can be susceptible to small changes in contextual variables. In other systems, advertising or manipulation of popularity could similarly disrupt the collective. Care must be taken in order to avoid undermining the simple mechanisms we have at our fingertips for promoting efficient accurate shared belief formation.

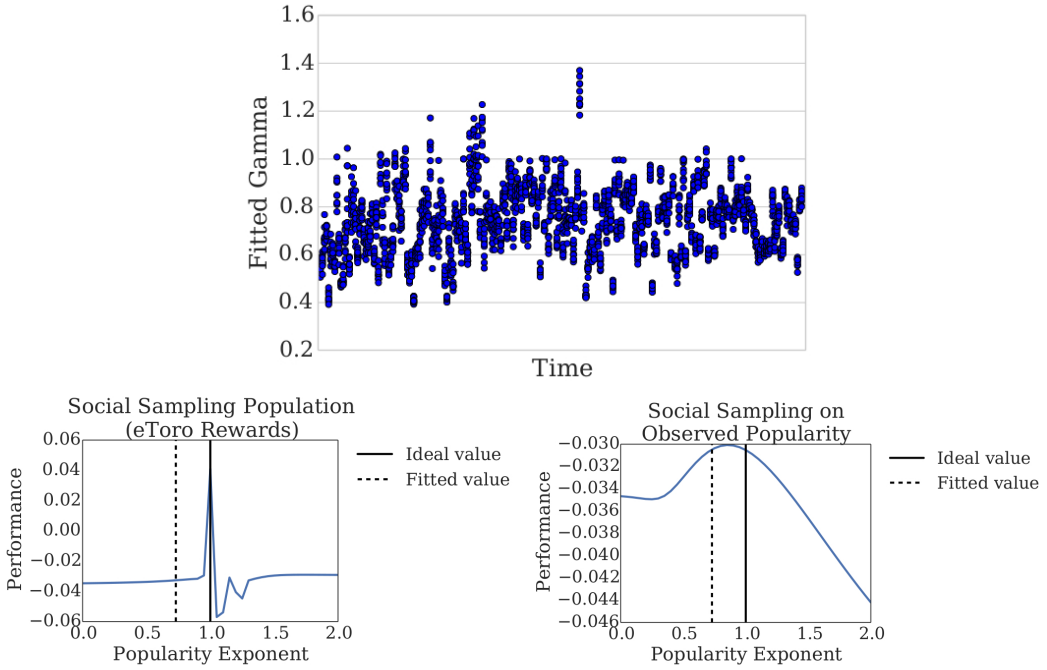

Figure 4-10: **Users on eToro sometimes achieve the optimal γ value but many times do not.** (Top) A plot γ fitted over time on each day, with bootstrap confidence intervals. (Left) Results from retrospective simulations of a population of social samplers on the eToro data. As in our idealized simulations, $\gamma = 1$ would be the optimal value for collective learning on eToro, assuming the entire population of followers is using social sampling. The value of γ from fitting to all the data simultaneously is shown. (Right) Results from simulations of social sampling given the observed popularity values on each day, rather than ideal values from a population of social samplers. When popularity is not generated by a population of individuals using social sampling with $\gamma = 1$, then a wider range of γ values achieves similar levels of performance.

4.9 Discussion

Human groups have an incredible capacity for technological, scientific, and cultural creativity. Our historical accomplishments and the opportunity of modern networked society to stimulate ever larger-scale collaboration have spurred broad interest in understanding the problem-solving abilities of groups—their collective intelligence. The phenomenon of collective intelligence has now been studied extensively across animal species [27]; collective intelligence has been argued to exist as a phenomenon distinct from individual intelligence in small human groups [126]; and the remarkable abilities of large human collectives have been extensively documented [110]. However, while the work in this area has cataloged what groups can do, and in some cases the mechanisms behind how they do it, we still lack a coherent formal perspective on what human collective intelligence actually is. There is a growing view of group behavior as implementing distributed algorithms [55, 61, 26, 32], which goes a step beyond the predominant analytical framework of agent-based models in that it formalizes specific information processing tasks that groups are solving. Yet this perspective provides little insight into one of the key features of human group cognition—the formation of shared beliefs [84, 18, 112, 99, 80, 101].

Recent work in cognitive science has provided a way to understand belief formation at the level of individual intelligence. One productive framework treats people as approximately Bayesian agents with rich mental models of the world [43, 41]. Beliefs that individuals hold are viewed as posterior distributions, or samples from these distributions, and the content and structure of those beliefs come from the structure of people's mental models as well as their objective observations. Belief formation is viewed as approximate Bayesian updating, conditioning these mental models on an individual's observations. We propose to model human collective intelligence as distributed Bayesian inference, and we present the first empirical evidence for such a model from a large behavioral dataset. Our model shows how shared beliefs of groups can be formed at the individual level through interactions with others and private boundedly rational Bayesian updating, while in aggregate implementing a rational

Bayesian inference procedure.

The more general utility of this framework will come from providing a principled constraint on individual-level mechanisms in modeling collective behavior, and from providing a rigorous way to relate these individual-level mechanisms to group-level models. Looking to the literatures on computer science, statistics, and signal processing for relevant distributed inference algorithms is likely to yield a fruitful path towards a new class of models of human collective behavior. At the same time, instances of this new class of models, such as our social sampling model, will bring novel algorithms to the area of distributed Bayesian inference.

4.10 Materials and Methods

4.10.1 Data Source

We received our data from a company called eToro. The data was generated from the normal activity of users of their website, etoro.com. The two main features of the eToro website during the time our dataset was being collected were a platform that allowed users to conduct individual trades and a platform for finding and following other users of the site. We will refer to the site's users interchangeably as either users or traders—having these two terms will ultimately reduce the ambiguity in some of our descriptions. The internal algorithms and the website design have changed over time, but the following description represents to the best of our knowledge the main contents and features of the website during the time period of our data.

The eToro website includes basic functionality for use as a simple trading platform. This platform allows users to enter long or short positions in a variety of assets. Entering a long position simply consists of buying a particular asset with a chosen currency. Entering a short position consists of borrowing the same asset to sell on the spot, with a promise to buy that asset at a later time. Taking a long position is profitable if the price of the asset increases, while taking a short position is profitable if the price of the asset decreases. Users can also enter leveraged positions. A leveraged

position is one in which an user borrows funds in order to multiply returns. Leveraged positions have more risk because users will lose their own investment at a faster rate if the price of the asset decreases.

At the time our data was collected, eToro focused on the foreign exchange market, so the trading activity mainly consisted of users trading in currency pairs—buying and selling one currency with another currency. However, users were also able to buy or sell other commodities such as gold, silver, and oil, and eventually certain stocks and bundled assets. The average amount of money invested in individual trades on eToro was about $30, and, after accounting for leverage, individual trades on average result in about $4000 of purchasing power. These amounts are small compared to the trillions of dollars traded daily in the foreign exchange and commodity markets[3], so individual traders are unlikely to have substantial market impact with their trades.

Besides providing a platform for individual trading, eToro also offers users the ability to view and mimic the trades of other users on their website. To be clear in our terminology, when referring to one user following another user to mimic their trades, we will call the first user the "following user" and the second the "target user". When referring to a specific mimicked trade, we will refer to the original trade as the "parent trade" and the copy as the "mimicked trade". eToro refers to "following" as "copying", and "mimickers" as "copiers". We use the term "follow" rather than "copy" so that we can reserve the word "copy" for social influence due to information about popularity, as in "copying the crowd in making decisions about whom to follow". eToro also offers an option to "follow" users without mimicking their trades, but we do not have data on these follow relationships.

While there is functionality for mimicking individual trades on eToro, we focus on the website's functionality for mirroring all the trades of specific users. Mirroring works as follows. First, a following user allocates funds that will be used for mirroring a target user. The following user's account then automatically executes all of the trades that the target user executes. The sizes of these trades are scaled up or down

[3]According to the Bank for International Settlements' 2013 "Triennial Central Bank Survey", the foreign exchange market (in which most of the trading on eToro occurs) has a daily trading volume of trillions of USD.

according to how much money the following user has allocated as funds for that follow relationship. When beginning a follow relationship, the following user can specify either to only mimic new trades of the target user or to also open positions that mirror all the target user's existing open positions. When a user stops following a target user, the following user can choose to either close all the open mimicked trades associated with that relationship or to keep those trades open.

There are certain limitations that eToro places on mimic trading. For example, users can follow no more than 20 target users with no more than 40% of available account funds allocated to a single target user. Users can also make certain adjustments to their mimicked trades. For example, following users can close a trade early or adjust a trade's "stop loss" amount.

eToro also offers an interface to assist users in finding traders to follow. The central feature of this interface at the time our data was collected was a tool that presented a list of other users on the site. This list could be sorted either by the number of followers those users had or by various performance metrics, such as percentage of profitable weeks or a metric called "gain". In a separate part of the site, users also had realtime or near realtime access to details of individual trades being executed by other users of the site.

In addition to searching for basic information using these tools, the website also allows users to view more detailed profiles of other traders on the site. These profiles present information such as the number of followers the user has had over time, the "gain" of the user over time, and information about opened and closed trades.

4.10.2 Data Processing

The dataset provided to us by eToro consists of a set of trades from the eToro website. The aggregated data we used in our analyses will be released upon publication. The entire dataset contains trades that were closed between June 1, 2011 and November 21, 2013. Our initial explorations (which multiple authors engaged in) must be assumed to have touched all of the data. When we began systematic analysis, in order to have a held-out confirmatory validation set for this analysis, we split the dataset into two

years. We use the first year of data for all our main analysis. We use the second year for a robustness check.

Each entry in the dataset we received includes a unique trade ID, a user ID, the open date of the trade, the close date of the trade, the names of the particular assets being traded, the amount of funds being invested, the number of units being purchased, the multiplying amount of leverage being used to obtain those units, the open rate of the pair of assets being traded, the close rate of that pair, and the net profit from the trade. For entries associated with mimickied trades, there is additional information. For individually mimicked trades, the parent trade ID is included. For trades resulting from follow relationships between users, "mirror IDs" are included in addition to parent trade IDs. A mirror ID is an integer that uniquely identifies a specific follow relationship. When a user begins to follow another user, a new mirror ID for that pair is created.

In order to study the relationship between previous popularity, perceived quality, and the follow decisions of users on eToro using this dataset, we first had to extract the popularity and the performance of each user on each day. For our main analysis, to best match the statistics that the eToro interface presented to users, we defined performance as investing performance measured as average return on investment (ROI) from closed trades over a 30-day period. In this computation, the ROI for a trade is determined by the profit generated from the trade divided by the amount withdrawn from the user's account to make the trade. If on a particular day a user did not make any trades in the previous 30 days, the performance of that user is not defined and the user is removed from the analysis for that day. This exclusion criterion removes inactive users from the analysis.

We use ROI rather than a risk-adjusted performance metric because the most prominent performance metric presented to users on eToro is what eToro calls "gain". eToro states this metric is computed using a type of "modified Dietz formula", an equation closely related to ROI. Thus ROI should better capture the perceived objective quality of following each user, though perhaps not true underlying quality.

We estimated the number of followers each user had on each day from the "mirror

IDs" present in the data we received. We first identify which two users are participating in each mirror ID, and we then determine the duration of the follow relationship between those two individuals as beginning on the first date we observe that mirror ID and ending on the last date we observe that mirror ID. From these time intervals we can then estimate the number of followers each user has on each day, as well as when each follow relationship begins.

We conduct our main analysis using approximately one year of data, from June 1, 2011 to June 30, 2012. However, we skip the first month of data when analyzing changes in popularity so that we are only using the period of time for which we have accurate estimates of previous popularity and changes in popularity. We also do not analyze changes in popularity that occur over weekends. Only a small percentage of trades occur on weekends since trading on eToro is closed on Saturdays and opens late on Sundays. Since the way we measure changes in popularity depends on having observed trades, days on which there is little to no trading can lead to inaccurate estimates of changes in popularity.

The data frames that we ultimately use for our statistical analyses, modeling, and model predictions are then constructed as follows. Each user is given a row in the data frame for each day on which that user had any trades in the previous 30 days. The columns associated with this row are the performance score of the user (the expected ROI from closed trades from the previous 30 days), the number of followers the user had on the previous day, the number of new followers that user gained on that day, and the number of followers that the user lost.

We also include two additional columns for each row, i.e. for each user-day pair, that include partially disaggregated trading information for each day. The first additional column is the actual amount of funds invested in new trades on that day by that user. This "actual amount" is the amount in USD invested in each trade initiated on that day after accounting for the leverage the trader used in those trades. The second additional column we use contains the sum of the realized and unrealized gains and losses from all new trades each user made on each day. To obtain the realized profit or loss for a user on a particular day, we simply sum the profit from each trade

that the user opened on that day and subsequently closed on the same day. To obtain the unrealized profit or loss, we take each trade that the user opened on that day but did not close. We then compute the unrealized profit or loss of those trades as the profit or loss that would have resulted from those trades being closed at the end of the day. To obtain the close rates for these trades, we use an external database of foreign exchange rates for all the currency pairs, and for other assets (whose close rates we don't have from an external database) we use the last observed rate on that day from the eToro data.

4.10.3 Imputing Missing Trades

There are a small number of observed mimicked trades that lack parent trades in the dataset. Since these missing parent trades are predominantly unprofitable, and since overestimating the performance of popular traders could substantially bias our results, we developed a method for recovering these missing parent trades. Certain fields of mimicked trades (including the open dates, close dates, assets traded, trade directions, and associated open and close rates) are typically almost identical to the fields of their associated parent trades. These similarities allow us to recover the direction of profit of these missing trades with high reliability. However, the initial amounts invested and the units purchased in missing parent trades are more difficult to infer. To estimate the amounts invested in each missing parent trade, we use the fact that the ratios between the units invested in mimicked trades and the units invested in their associated parent trades are relatively stable for a particular mirror ID. Specifically, we use the following procedure. We first, for each mirror ID, compute the median ratio between the units invested in each of the observed mimicked trades associated with that mirror ID and the units in each of those trades' parents. For a particular missing parent, we then find all of the mimicked trades in our data with this missing trade as a parent. We then gather what the units invested in the parent trade would be according to each of the median ratios associated with those mimicked trades, and we take the median of those unit values. We use this final quantity as the number of units we presume were purchased in the original trade. We finally

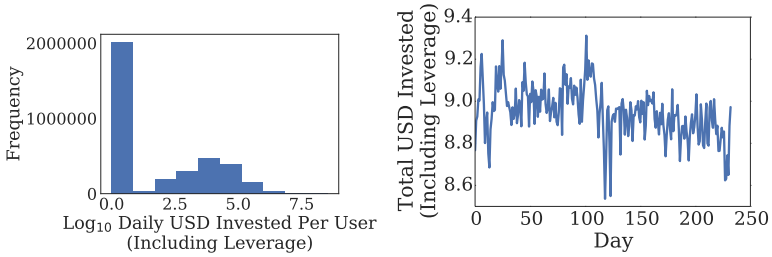

Figure 4-11: Statistics on the amount invested by all users on eToro per day.

compute the amount of funds invested in each missing trade and the ultimate profit made from each of those trades from the inferred units purchased and the open and close rates of a single observed mimicked trade. We conduct our main analysis using these imputed trades.

4.10.4 Descriptive Statistics

We now provide descriptive statistics of the eToro data in our main analysis. Figure 4-11 indicates that many users are sporadic in their investment activities, but the total amount of money invested per day on eToro remains fairly stable over time. Figure 4-12 indicates that mean ROI on eToro varies around zero but tends to be slightly negative and appears to have a slightly increasing trend over time. Figure 4-13 indicates that the total number of followers on eToro increases over time, but the maximum popularity on a particular day divided by the total number of followers on that day remains relatively stable over time. Figure 4-14 shows that the vast majority of users have zero followers, and the distribution of followers is long-tailed.

4.10.5 Regression Analysis

We now provide statistical evidence that there is a positive multiplicative interaction between previous popularity and performance in relation to future popularity. To support this hypothesis we perform both a simple analysis using ordinary linear models and a more robust analysis that accounts for dependence between data points and individual user-level effects. Our model comparisons described in the subsequent sections provide further evidence for this interaction effect. For the following analy-

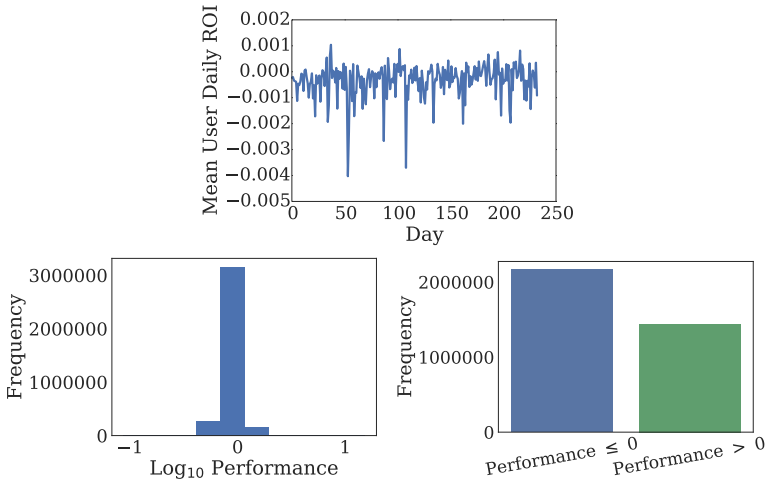

Figure 4-12: Statistics on user return on investment.

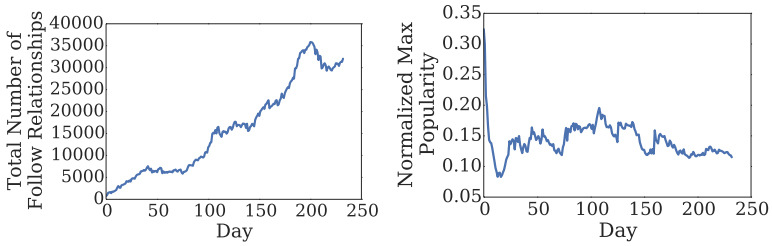

Figure 4-13: Statistics on the number of followers and maximum popularity over time.

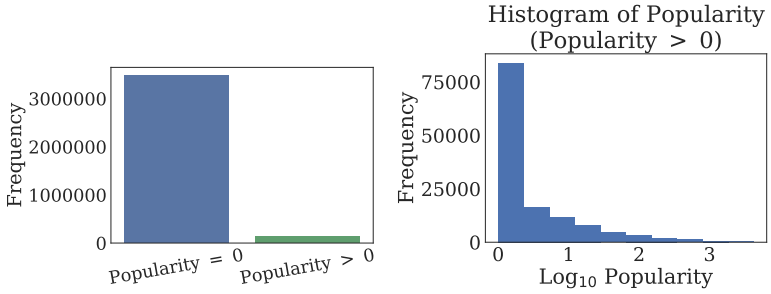

Figure 4-14: Statistics of user popularity.

ses, we use the data format as described at the end of the "Data Processing" section above, and we use two-sided hypothesis tests for all p-values.

Table 4.1: Results from an ordinary linear regression.

Independent Variable	**Coefficient**	p**-value**
Intercept	8.226e-03	<2e-16
Popularity	6.499e-03	<2e-16
Performance	1.720e-02	3.01e-06
Interaction	4.748e-02	<2e-16

Initially, we perform a basic statistical analysis assuming that the observed changes in popularity within each user and within each day are conditionally independent of each other given previous popularity and performance. That is, this first analysis assumes that the only pieces of information that affects follow decisions are popularity and performance information, and there are no systematic preferences for particular users and no systematic differences in changes in popularity on particular days. The results of this ordinary linear regression using all of the active users on each day are given in Table 4.1. The dependent variable in this regression is the change in popularity of each user on each day computed as the user's popularity on that day minus the user's popularity on the previous day. The independent variables are the performance scores of each user computed from the previous 30 days, the number of followers that the user had on the previous day, and an interaction term multiplying these two values. The results of this regression indicate that there is a significant positive interaction effect between previous popularity and previous performance in relation to changes in popularity.

Table 4.2: Results from fixed effects model with robust standard errors.

Independent Variable	**Coefficient**	p**-value**
Popularity	1.3301e-03	0.4007003
Performance	1.0583e-02	0.1814710
Interaction	4.3830e-02	0.0115439

For a more robust statistical analysis, we used a fixed effects model with Arellano robust standard errors [5]. We included fixed effects for users and days. Including fixed effects for users and days allows us to rule out the possibility that the effects

we observe are driven by a few peculiar users or days, and the robust standard errors correct the p-value for non-equal variance and correlated error terms. Besides these changes, the models are the same as the previous ordinary linear models. The results of these fixed effect models are given in Table 4.2.

4.10.6 Predicting Follow Decisions

Our model comparison analysis relied on computing the follow decisions users on eToro would have been expected to have made according to the social sampling model and alternative models. In all cases we examine aggregations of these decisions in the form of predicting the total number of new followers each trader on eToro obtains. More specifically, we predict the number of new followers each user gets on each day given the performance and popularity of that user (and of every other user) on the previous day, and given the total number of new follow events we observe on those days. To get these aggregates, we predict individual follow decisions, which are constrained by the traders each user is already following.

In the social sampling model, decision-makers make decisions independently, so the probability that a given decision-maker i chooses a specific new option j at time t is given by the decision probability ϕ_{jt}^{main},

$$\phi_{ijt}^{main} = \frac{\eta^{x_{jt}}(1-\eta)^{1-x_{jt}} \cdot (p_{jt}+\epsilon_t)^{\gamma}}{\sum_{k \notin followed_{it}} \eta^{x_{kt}}(1-\eta)^{1-x_{kt}} \cdot (p_{kt}+\epsilon_t)^{\gamma}},$$

where $\epsilon_t > 0$ is a small smoothing parameter that ensures all users have some probability of gaining followers, x_{jt} in the case of the eToro data simply indicates whether user j has positive or negative performance on day t, and followed_{it} tracks the traders user i is currently following at time t (since a trader cannot be followed twice). We arbitrarily choose $\epsilon_t = \frac{1}{M_t}$, where M_t is the number of active users on day t. Given $\phi_{ijt}^{main} = 0$ if $j \in followed_{it}$, the expected number of new followers user j gets at time t is then $\sum_{i \notin followers_{jt}} \phi_{ijt}^{main}$, where $followers_{jt} = \{i \,|\, j \notin followed_{it}\}$ is the set of followers of j at time t.

4.10.7 Alternative Models

We consider a set of alternative models in order to identify how well the social sampling model is able to account for structure in the follow decisions present in our data compared to alternative plausible models. We specify these alternative models in terms of "decision probabilities" analogous to ϕ_{ijt}^{main}. These decision probabilities provide the probability under each model that an individual will decide to commit to a particular option (or follow a particular trader in the case of eToro). It is possible to specify all our alternative models in this way because every model we consider assumes that all decisions are conditionally independent given the popularity and performance of every option.

The first of these alternatives is a proxy for a probability matching rational agent model that we call the "Performance Regression" model. This model uses only performance information, and does not reduce the performance signals to being binary as our social sampling model does. The decision probability under this model is

$$\phi_{ijt}^{perf} \propto \sigma(\beta_0 + \beta_1 q_{jt}),$$

where σ is the logistic function, the β_i variables are free parameters, and q_{jt} is a real-valued performance signal (30-day rolling average daily ROI on closed trades for our main analysis). The performance regression allows us to evaluate the predictive power of using performance information alone to predict follow decisions.

The next alternative we consider is an extended regression model that we call the "Full Regression" model. This alternative consists of a generalized linear model that includes an interaction term between popularity and performance. Such a model could conceivably generate the multiplicative interaction effects we observe in the eToro data, but lacks some of the additional structure that the social sampling model has. The full regression assumes that a decision-maker chooses to commit to option j with probability

$$\phi_{ijt}^{full} \propto \sigma(\beta_0 + \beta_1 q_{jt} + \beta_2 p_{jt} + \beta_3 q_{jt} p_{jt}),$$

where the notation is as above. The purpose of comparing to this alternative is to test whether having the additional structure of including popularity as a prior in the social sampling model lends additional predictive power, as compared to having a heuristic combination of popularity and performance.

We also consider a reduction of the social sampling model that does not use performance information. We call this model the "Popularity Model". Under the popularity model the decision probability becomes

$$\phi_{jt}^{pop} \propto (p_{jt} + \epsilon_t)^{\gamma},$$

and again we use $\epsilon_t = \frac{1}{M_t}$ as the smoothing parameter. Comparing to this preferential attachment model allows us to understand how much predictive power we get from including performance information while controlling for the structure of how social information is used in the social sampling model. This preferential attachment model is a canonical simple heuristic model of social decision-making.

We also consider an alternative model that is a reduction of the social sampling model that uses only performance information. This model, which we refer to as the "Performance Model", uses the decision probability

$$\phi_{ijt}^{perfm} \propto \eta^{x_{jt}}(1-\eta)^{1-x_{jt}}.$$

Since the performance model is the one that would be obtained from the social sampling model when all options have the same popularity, this model allows us to predict how decision-makers might behave if they did not have social information.

Our final alternative model is an additive combination of the popularity model and the performance model. This model, which we call the "Additive Model", represents a situation in which some agents choose whom to follow based on preferential attachment while other choose based on performance. Under this model the decision probability becomes

$$\phi_{ijt}^{add} \propto \alpha\phi_{ijt}^{pop} + (1-\alpha)\phi_{ijt}^{perfm},$$

where $\alpha \in [0, 1]$ is a free parameter and again we use $\epsilon_t = \frac{1}{M_t}$ as the smoothing parameter. Comparing to the additive model allows us to verify that popularity and performance are combined multiplicatively rather than additively.

4.10.8 Parameter Fitting

To estimate the parameters of these models we use a maximum likelihood procedure. Letting T denote the number of days we observe, letting F_t denote the total of new follow decisions on day t, letting $source_{ft}$ denote the index of the follower in follow decision f on on day t, letting $target_{ft}$ denote the index of the trader followed in follow decision f on on day t, the likelihood of the parameters given all of the new follow decisions is

$$\prod_{t=1}^{T} \prod_{f=1}^{F_t} \phi_{source_{ft}, target_{ft}, t},$$

where the ϕ is determined by whichever of the models we are fitting. We then use numerical optimization to maximize this likelihood and obtain the best fitting α, β_i, γ, and η parameters. When we fit the parameters of the social sampling model over time, we simply compute the likelihood function restricted to the data points for each day individually.

4.10.9 Model Fits

To quantitatively compare the fit of social sampling with our alternative models, we compare the mean squared prediction error (MSE) of each model. Prediction error is given by the difference between the expected number of new followers each trader obtains and the actual number. The eToro dataset is highly unbalanced, with most users having zero followers. To adjust for this imbalance, in addition to using marginal mean squared error we also look a weighted mean squared prediction error, with weights given by the inverse of the probability of sampling each data point. We estimate the probability of sampling a data point by binning the popularity and performance of each user on each day (according to the bins used in Figure 4-3), and counting the number of user-day pairs that fall into each bin.

Table 4.3: Results of model comparison.

Model	**Unadjusted MSE**	**Adjusted MSE**
Popularity Model	2.18805045175	618.976488667
Additive Model	2.11109662634	636.617485337
Social Sampling	2.15097403109	596.668627703
Performance Regression	3.07280958801	923.112970477
Full Regression	2.66340190835	767.600539636
Performance Model	3.07293303493	923.090627025

4.10.10 Robustness Checks

We also examine the adjusted MSE model fits using alternative performance metrics, and on our confirmatory dataset. We look shorter and longer windows for the rolling averages (7-day rolling average daily ROI and 60-day rolling average daily ROI), we test whether we obtain similar results when we do not deal with missing data, and we fit the social sampling model and the alternative models on our confirmatory dataset. Social sampling obtains the best fit on the longer rolling average window, but obtains slightly worse fit than the popularity model and the additive model on the shorter rolling average window. A reason for this difference could be that the statistics presented to users were 30-day, 60-day, and 90-day aggregates, so we would expect to see decisions being based on these longer time windows. Social sampling also obtains the best fit when we do not deal with missing data. On our confirmatory set with the 30-day ROI performance metric, social sampling is less competitive with the popularity model and the additive model. Social sampling still captures the multiplicative interaction effect between popularity and performance better on the confirmatory set, but the has poor overall performance due to social sampling underestimating the positive effect of high popularity in the second year of our data. It is possible that changes in the interface led to different behavior in this second year, or perhaps that the performance statistics eToro used were changed so that ours are no longer a good proxy, as with the shorter rolling average window results.

4.10.11 Consistency Checks

We executed two tests to provide further evidence for the specific parametric form of the social sampling model. We first examined the inferred value of the η parameter in the social sampling model. η represents the expected reward from committing to the best option, or in the case of the eToro data, the skill of the best trader as measured by the expected proportion of performance signals that will be greater than zero. To arrive at plausible actual values for what the skill of the "best trader" on eToro might be, we first rank all traders according to the amount of evidence for their success. For this ranking we use an aggregated single-day net profit values. We achieve the ranking of the traders by taking the total number of days each trader had positive single-day profit and subtracting the total number of days those traders had negative single-day profit. This metric simultaneously considers both the total amount of positive or negative evidence for trader skill in addition to the proportion of positive evidence. For each of the top ten traders according to this metric, we then compute a 95% Bayesian confidence interval (under a uniform prior) of the probability that those users will achieve positive performance, assuming a Bernoulli model. These confidence intervals are then plotted in Figure 4-7 along with the actual inferred "best skill" η parameter.

We also directly test a key assumption of the social sampling model. An idealized social sampling model predicts that the popularity-normalized number of new followers each trader is expected to obtain should be constant given performance. To check this assumption, we first calculate the number of new followers each trader obtains on each day, divided by those traders' popularity values on those days. We then compute a logarithmic binning of trader popularity and our usual binary binning our performance, and we plot the average popularity-normalized number of followers in each bin. These values are plotted in Figure **??**. The predicted constant value of popularity-normalized new followers for high performing individuals is obtained by taking the average number for poor performing individuals and multiplying that number by $\frac{\eta}{1-\eta}$, as would be the case in the social sampling model in a fixed-size

population.

4.10.12 Normative Posterior

To assess the macro-level predictions of the social sampling model we compute

$$P(j \text{ is good} \,|\, \boldsymbol{x}_{j,\leq t}) = \frac{(\eta^*)^{w_{jt}}(1-\eta^*)^{l_{jt}}}{(\eta^*)^{w_{jt}}(1-\eta^*)^{l_{jt}} + 0.5^{w_{jt}+l_{jt}}},$$

the probability each trader is good. Here, $w_{jt} = \sum_{d=1}^{d\leq t} \mathbb{1}(x_{jd} = 1)$ is the number of positive signals a trader has, and $l_{jt} = \sum_{d=1}^{d\leq t} \mathbb{1}(x_{jd} = 0)$ is the number of negative signals a trader has. We then normalize these values to obtain the probability given by Thompson sampling on this posterior: $\frac{P(j \text{ is good} \,|\, \boldsymbol{x}_{j,\leq t})}{\sum_k P(k \text{ is good} \,|\, \boldsymbol{x}_{k,\leq t})}$.

4.10.13 Idealized Simulations

To perform our idealized simulation, we implement the environment assumed in a theoretical justification of our model. In these simulations, there are M options that N agents can choose to commit to on each of T steps. Each of these options generates a reward, either 0 or 1 on each round, with the rewards chosen according to independent Bernoulli draws. We suppose that committing to a decision has a cost of 0.5. Then $M-1$ of these options have expected return 0, while one option—the "best option"—has positive expected return. In our simulations the best option has Bernoulli parameter η^*, which can be different from the agents' assumed η. At time t, the agents are able to observe the decisions made in round $t-1$, as well as the reward signals from round $t-1$. The agents in these simulations make their decisions according to the social sampling model strategy either with $\gamma = 1$ or with alternative scalings on popularity. In each round, every agent first selects an option to consider with probability proportional to $p_{jt}^{\gamma} + \frac{1}{M}$, where the notation is as above. Each agent then chooses to commit to the option that agent is considering with probability $\eta^{x_{jt}}(1-\eta)^{1-x_{jt}}$. When $\gamma = 0$ this process becomes the performance model. When $\eta = 0.5$ this process becomes the popularity model.

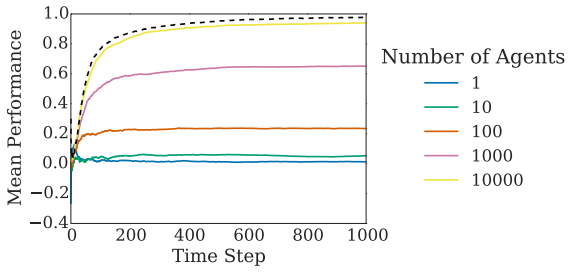

Figure 4-15: Performance of social sampling as a function of the number of agents in the population.

We conduct several sets of simulations. The first set examined the impact of alternative γ scaling exponents. For this experiment we look at the average reward in the final round achieved by agents who committed to some option in that round, which we call the "Mean Follower Performance". The results of these simulations are shown in Figure 4-7. Each line in this figure represents a different combination of simulation parameters. We look at all combinations of $N \in [1000, 5000, 10000]$, $M \in [5, 10, 100]$, $T \in [100, 500, 1000]$, $\eta^* \in [0.6, 0.7, 0.8]$. For this simulation experiment, we assume $\eta = \eta^*$. Each data point is an average over 500 repetitions for that particular combination of simulation parameters. The panels of the figure are separated by η^* value. Line color also indicates η^*, line size indicates N, line type indicates M, and transparency indicates T.

We also compare convergence over time across single-agent Thompson sampling, social sampling (with $\gamma = 1$), and group Thompson sampling. Here we fix the number of options to 100, and look at the case when one option has $\eta^* = 0.8$. We then plot the mean performance of each agent using each algorithm at each time step, and observe how this mean performance varies over time. The impact of varying the number of agents in this simulation in shown in Figure 4-15. Social sampling performs most consistently best when the number of agents is large compared to the number of options.

To conduct our evolutionary simulations, we simulate a number of generations of an infinite population. This population begins with 1% social samplers, and the rest independent decision-makers (social samplers with $\gamma = 0$. Each generation is

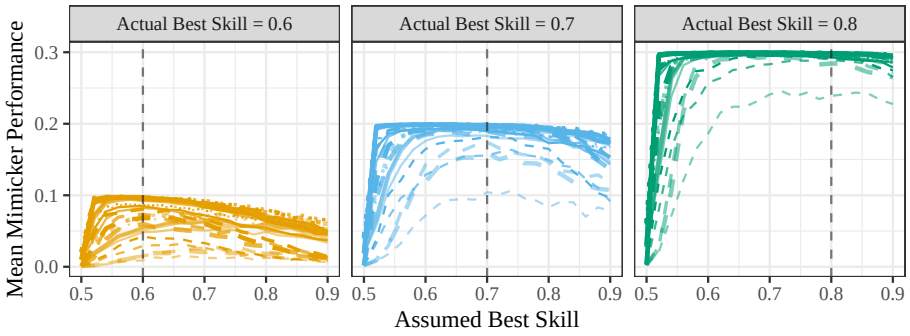

Figure 4-16: Performance of social sampling ($\gamma = 1$) with η values that do not match the idealized environment.

faced with 10 options and 50 time steps to learn about those options. A different option each generation has positive expected value, while the other options in each generation have equal change of positive or negative signals. The percentage of social samplers in each following generation is chosen according to the relative performance of the group of social samplers versus the group of independent decision-makers.

We also test how well the social sampling model performs in a setting where the environment model differs from the model assumed by the agents in our Bayesian justification. We again look at mean mimicker performance in the final round of each simulation. Here, though, we fix $\gamma = 1$ and consider η values ranging from 0.5 to 0.9, independent of the environment value of η^*. (Note, $\eta = 0.5$ corresponds to ignoring performance information.) The results of these simulations are shown in Figure 4-16. Each line in this figure represents a different combination of simulation parameters, and the parameter sweep is over the same space as our first set of simulations. The panels and line characteristics are also determined as in the first set of simulations. These simulations indicate that social sampling is relatively robust to deviations between agent's assumed environment model and the actual environment model.

We also report on a final set of simulations here to examine the extent to which a population can be manipulated by an initial artificial boost in popularity of an inferior option. In these simulations, rather than all options starting out with popularity $p_{j0} = 0$, one option $j' \neq j^*$ is initialized with $p_{j'0} = b$, where b represents a temporary

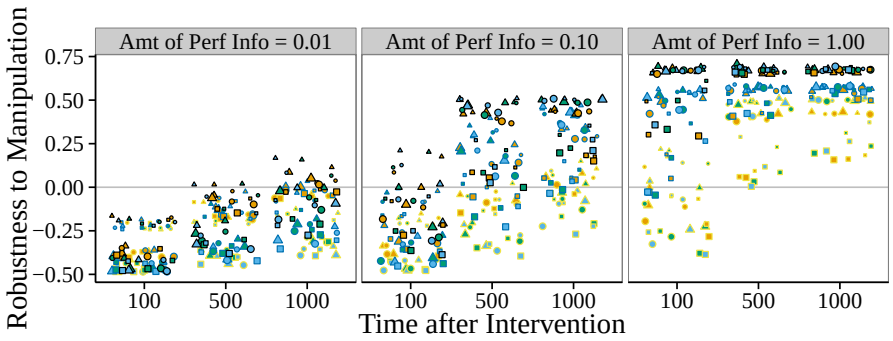

Figure 4-17: Simulated manipulations of social sampling.

"bump" in popularity. This bump in popularity is removed after the first round. In these simulations we also introduce another parameter of the simulations, which we call the "data rate". The data rate d represents the amount of performance information agents get about each option. In the simulations, the reward of each option in each round is hidden with probability d. If an agent samples an option to consider based on popularity that then turns out to have a hidden reward, the agent simply chooses to commit to the option with probability 0.5. As the outcome measure in these simulations we examine the difference in popularity in the final round between the best option and the "bumped" inferior option, while normalizing by the total number of agents, i.e. $\frac{p_{j^*T} - p_{j'T}}{N}$.

In this final set of simulations we take $\gamma = 1$. We also consider $b \in [1, 10, 100]$ and $d \in [0.01, 0.1, 1]$ in addition to the same parameter sweep as in the first set of simulations. The results of these simulations are shown in Figure 4-17. Each point in this figure is an average over 500 repetitions of the combination of simulation parameters the point represents. The border color of the points indicates η^*, the fill color indicates N, the shape indicates M, the transparency indicates T, the size indicates b, and the panel indicates d. Social sampling is robust to manipulation when enough data is available, but fragile otherwise.

4.10.14 Simulating Performance on eToro Data

In addition to our idealized simulations, we also retrospectively simulate the performance of social sampling using the actual profits and losses from trades on eToro. We do these simulations in two ways. The first way does not use popularity information as it is observed in the behavioral data at all. For this simulation, we compute what the expected popularity of each trader would be if an infinite population of social sampling agents was choosing between them, given the actual performance signals of those traders. These social sampling agents then use this simulated popularity for their own decisions. The second way uses the observed popularity values from eToro instead of ideal popularity values. We can think about the first simulation as the performance of a population of social samplers in the eToro environment, and we can think about the second type of simulation as what a single social sampler using the social information from the actual followers on eToro.

Chapter 5

Shared Belief Formation for Coordination

In the prior two chapters I have investigated mechanisms that promote accurate shared belief formation about the external world. A final modeling challenge I take up in this thesis is how a mechanism for forming accurate interpersonal shared belief contributes to coordination—solving the problem of aligned actions [65]. Whether in groups of humans or groups of computer agents, collaboration is most effective between individuals who have the ability to coordinate on a joint strategy for collective action. However, in general a rational actor will only intend to coordinate if that actor believes the other group members have the same intention. This circular dependence makes rational coordination difficult in uncertain environments if communication between actors is unreliable and no prior agreements have been made. An important normative question with regard to coordination in these ad hoc settings is therefore how one can come to believe that other actors will coordinate, and with regard to systems involving humans, an important empirical question is how humans arrive at these expectations. I introduce an exact algorithm for computing the infinitely recursive hierarchy of graded beliefs required for rational coordination in uncertain environments, and I introduce a novel mechanism for multiagent coordination that uses it. This algorithm is valid in any environment with a finite state space, and extensions to certain countably infinite state spaces are likely possible. I test this

mechanism for multiagent coordination as a model for human decisions in a simple coordination game using existing experimental data. I then explore via simulations whether modeling humans in this way may improve human-agent collaboration.

5.1 Introduction

Forming shared plans that support mutually beneficial behavior within a group is central to collaborative social interaction and collective intelligence [44]. Indeed, many common organizational practices are designed to facilitate shared knowledge of the structure and goals of organizations, as well as mutual recognition of the roles that individuals in the organizations play. Once teams become physically separated and responsiveness or frequency of communication declines, the challenge of forming shared plans increases. Part of this difficulty is fundamentally computational. In theory, coming to a fully mutually recognized agreement on even a simple action plan among two choices can be literally impossible if communication is even mildly unreliable, even if an arbitrary amount of communication is allowed [45, 76].

This problem is well-studied within the AI literature (e.g., [37]), though the core difficulties still manifest in contemporary research on "ad hoc coordination"—collaborative multiagent planning with previously unknown teammates [108]. However, surprisingly little is known about the strategies that *humans* use to overcome the difficulties of coordination [113]. Understanding how and when people try to coordinate is critical to furthering our understanding of human group behavior, as well as to the design of agents for human-agent collectives [56]. Existing attempts at modeling human coordination have focused either on unstructured predictive models (e.g., [33]) or bounded depth socially recursive reasoning models (e.g., [34, 127]), but there is reason to believe that these accounts miss important aspects of human coordination.

One concept that appears repeatedly in formal treatments of coordination but has not appeared meaningfully in empirical modeling is common knowledge. Two agents have common knowledge if both agents have infinitely nested knowledge of

the other agent's knowledge of a proposition, i.e. the first agent knows the second agent knows, the first agent knows the second agent knows the first agent knows, etc. Common knowledge has been shown to be necessary for exact coordination [45], and a probabilistic generalization of common knowledge, called common p-belief, has been been shown to be necessary for approximate coordination [87]. While these notions are clearly important normatively, it is not entirely clear how important they are empirically in human coordination. Indeed, supposing that humans are able to mentally represent an infinitely recursive belief state seems a priori implausible, and the need to represent and infer this infinite recursive belief state has also been a barrier to empirically testing models involving common knowledge.

Nevertheless, building on the existing normative results, a group of researchers recently designed a set of experiments to test whether people are able to recognize situations in which common knowledge might obtain [113] (hereafter referred to as the "Thomas experiments"). These researchers argued that people do possess a distinct mental representation of common knowledge by showing that people will attempt to coordinate more often in situations where common knowledge can be inferred. However, this previous work did not formalize this claim in a model or rigorously test it against plausible alternative computational models of coordination. This existing empirical work therefore leaves open several important scientific questions that a modeling effort can help address. In particular: How might people mentally represent common p-belief? Do people reason about graded levels of common p-belief, or just "sufficiently high" common p-belief? Finally, what computational processes could people use to infer common p-belief?

In this work I use a previously established fixed point characterization of common p-belief [87] to formulate a novel model of human coordination. In finite state spaces this characterization yields an exact finite representation of common p-belief, which I use to develop an efficient algorithm for computing common p-belief. This algorithm allows me to simulate models that rely on common p-belief. Because of the normative importance of common p-belief in coordination problems, this algorithm may also be independently useful for coordination in artificial multiagent systems. I show using

data from the Thomas experiments that this model provides a better account of human decisions than three alternative models in a simple coordination task. Finally, I show via simulations based on the data from the Thomas experiments that modeling humans in this way may improve human-agent coordination.

5.2 Background

We first provide a description of the coordination task we will study in this chapter: the well-known coordinated attack problem. We then provide an overview of the formal definitions of common knowledge and common p-belief, and their relationship to the coordinated attack problem.

5.2.1 Coordinated Attack Problem

The coordination task that we study in this chapter is alternatively called the coordinated attack problem, the two generals problem, or the email game. The original formulation of this task was posed in the literature on distributed computer systems to illustrate the impossibility of achieving consensus among distributed computer processors that use an unreliable message-passing system [45], and the problem was later adapted by economists to a game theoretic context [100]. Here we focus on the game theoretic adaptation, as this formulation is more amenable to decision-theoretic modeling and thus more relevant for modeling human behavior.[1]

In this task the world can be in one of two states, $x = 1$ or $x = 0$. The state of the world determines which of two games two players will play together. The payoff matrices for these two games are as follows ($a > c > \max(b, d)$):

$x = 1$	A	B
A	a,a	b,c
B	c,b	c,c

$x = 0$	A	B
A	d,d	b,c
B	c,b	c,c

[1] Our exposition largely assumes familiarity with rudimentary game theory, and familiarity with measure-theoretic probability as it appears in incomplete information games.

The players receive the optimal payoff if they coordinate on both playing A when $x = 1$, but playing A is risky. Playing A is inferior to playing B if $x = 0$ or if there is a mismatch between the players' actions. Playing B is safe with a sure payoff of c. Thus in order for it to be rational to play A, a player must believe with sufficient confidence both that the world state is $x = 1$ and that the other player will play A.

5.2.2 Common p-Belief

In order for a player to believe the other player will play A in this game, it is not enough for that player to believe that the other player knows $x = 1$. If the second player does not believe that the first player knows $x = 1$, then the second player will not try to coordinate. Therefore the first player must also at least believe that the second player believes the first player knows $x = 1$. However, it turns out even this amount of knowledge does not suffice. In fact, an infinite hierarchy of recursive belief is needed [88]. This infinite hierarchy of beliefs has been formalized using a construct called common p-belief, which we now define.

Using standard definitions from game theory, we define a two-player finite Bayesian game to be a tuple $((\Omega, \mu, (\Pi_0, \Pi_1)), (\mathcal{A}_0, \mathcal{A}_1), (u_0, u_1))$ consisting of a finite state space $\Omega = \{\omega_1, \ldots, \omega_{|\Omega|}\}$, a probability measure μ defined over that state space, the information partition Π_i of each player i, the action set $\mathcal{A}_i$ of each player, and the utility function u_i of each player. Elements of Ω are called states, and subsets of Ω are called events. For a given world state ω and player i the partition of Ω, Π_i, uniquely specifies the beliefs of player i in the form of posterior probabilities. $\Pi_i(\omega)$, which indicates the unique element of Π_i containing ω, can be thought of as the observation that player i receives when the true state ω occurs. Specifically for any event $E \subseteq \Omega$, $\mu(E \,|\, \Pi_i(\omega))$ is the probability that player i assigns to E having occurred given ω has occurred. As a shorthand, we write $P_i(E \,|\, \omega) = \mu(E \,|\, \Pi_i(\omega))$. Using another common shorthand, we will treat propositions and events satisfying those propositions interchangeably. For example, in the coordinated attack problem we will be interested in whether there is "common p-belief" that $x = 1$, which will refer to common p-belief in the event $C = \{\omega \in \Omega : x(\omega) = 1\}$, where x formally is a random variable mapping

$\Omega \to \{0, 1\}$.

Following [28], we say that player i p-believes[2] an event E at ω if $P_i(E \,|\, \omega) \geq p$. An event E is said to be p-evident if for all $\omega \in E$ and for all players i, player i p-believes E at ω. In a slight divergence from the standard terminology of this literature, we say an event E is super-p-evident if for all $\omega \in E$ and for all players i, $P_i(E \,|\, \omega) > p$ (the only difference being strict inequality). We say there is common p-belief in an event C at state ω if there exists a p-evident event E with $\omega \in E$, and for all $\omega' \in E$ and all players i, player i p-believes C at ω'. Common knowledge is defined as common p-belief for $p = 1$.

A critically important result of [87] states that this definition of common p-belief is equivalent to a more intuitive infinitely recursive formulation. The importance of this definition of common p-belief is therefore that it provides a fixed point characterization of common p-belief strictly in terms of beliefs about events rather than directly in terms of beliefs about other players. When Ω is finite, common p-belief can thus be represented in terms of a finite set of states, rather than an infinite hierarchy of beliefs.

5.3 Models

We now describe four strategies for coordination in the coordinated attack game we study. Two of these strategies involve the computation of p-evident events and common p-belief, which we will use to test whether human coordination behavior could be explained in terms of reasoning about p-evident events. The other two strategies serve as baselines.

5.3.1 Rational p-Belief

The first strategy we consider represents an agent who maximizes expected utility at an equilibrium point of the coordinated attack problem we study. The strategy

[2] We use an italicized "p" when referring to specific values of p-belief and a non-italicized "p" when referring to the terms in general.

is implemented as follows: player i plays action A if and only if the player believes with probability at least $p^* = \frac{c-b}{a-b}$ that both players have common p^*-belief that $x = 1$. This strategy forms an equilibrium of the coordinated attack problem if $p^* > P_i(x = 1)$ and if evidence that $x = 1$ always leads to certain belief that $x = 1$, i.e. $P_i(x = 1 \,|\, \omega) > P_i(x = 1) \Rightarrow P_i(x = 1 \,|\, \omega) = 1$ for all ω. These conditions will be satisfied by the specific state spaces and payoffs we use to represent the Thomas experiments. We call this model the rational p-belief strategy. The proof that this strategy forms an equilibrium, including a derivation for the specific form of p^*, is included in our supplementary materials.

5.3.2 Matched p-Belief

The second strategy we consider is a novel probabilistic relaxation of the rational p-belief strategy. Humans have been shown to exhibit a behavior called probability matching in many decision-making settings [49, 121]. Probability matching consists of taking the probability p that a decision is the best decision available, and choosing to make that decision with probability p. While probability matching is not utility maximizing, it can be viewed as rational if players are performing sample-based Bayesian inference and if taking samples is costly [120]. Motivated by this frequently observed behavior, we propose a model we call the matched p-belief strategy. A player i using this strategy chooses action A at ω with probability p equal to the maximal common p-belief that the player perceives at ω, i.e. the largest value such that i p-believes at ω that there is common p-belief that $x = 1$.

5.3.3 Iterated Maximization

Next we consider a well-known model of boundedly rational behavior sometimes called a "level-k" depth of reasoning model. This family of models has been shown to be consistent with human behavior in a diversity of settings, including some coordination games (e.g., [127]), and hence is a strong baseline. Since the term "level-k" is used for many slightly different models, we call our instantiation of this model the iterated

maximization strategy. This strategy assumes that players have a certain fixed level of recursive social reasoning k. A player using the level-k iterated maximization strategy chooses the action that maximizes that player's expected utility when playing with a player using the level-$(k-1)$ strategy. The level-0 player takes action A at ω if $P_i(x = 1 \,|\, \omega) > \frac{c-b}{a-b}$. This level-0 strategy corresponds to the player maximizing expected utility assuming the player can control the actions of both players, or equivalently that the optimal joint action is taken according to that player's beliefs. While in general the predictions of level-k models depend strongly on the specification of the level-0 strategy, in informal exploration we found that the qualitative conclusions of our work are robust to whether we instead specify the level-0 strategy as always playing A or choosing between A and B uniformly randomly.

5.3.4 Iterated Matching

Finally, we also consider a less common depth of reasoning model that combines the iterated maximization strategy with probability matching behavior, which we call iterated matching. Like the iterated maximization strategy, this strategy assumes that players have a certain fixed level of recursive social reasoning k. However, instead of choosing the action that maximizes expected utility, a level-k player using the iterated matching strategy chooses to take action A with probability equal to that player's belief that $x = 1$, times the expected probability that a level-$(k-1)$ companion player would play A. The level-0 player probability matches on $P_i(x = 1 \,|\, \omega)$.

5.4 Algorithms

In this section we present the algorithms we use to implement each of the models we consider. To the best of our knowledge the existing literature on common p-belief has yet to offer algorithms for computing common p-belief (or in our case the perceived maximal common p-belief) for a given world state and observation model. This computation is central to the rational and matched p-belief strategies. Hence we offer the first fully computational account of coordination via reasoning about p-evident

events. Algorithms for iterated reasoning are straightforward and well-known.

The challenge in developing an algorithm for computing a player's perception of the maximal common p-belief is avoiding enumeration over all exponentially many possible subsets of Ω. While it is straightforward to evaluate whether a particular given event is p-evident, the definition of common p-belief requires only the existence of some such event. Computing perceived maximal common p-belief therefore requires jointly searching over values of p and over subsets of Ω. We leverage the generic mathematical structure of p-evident events in finite state spaces in order to develop an exact algorithm that avoids enumerating all subsets. Our algorithm only requires a search that is polynomial in the size of the state space. Of course, the state spaces in many problems are often themselves exponential in some underlying parameter variables, and hence future improvements on this algorithm would be desirable. Extensions to at least certain countably infinite or continuous state spaces are likely possible as well, such as perhaps by refining the search to only consider events that have non-zero probability given the player's observations.

5.4.1 Computing Information Partitions

Our algorithms require access to the information partitions of each player. However, directly specifying the information partitions that people have in a naturalistic setting, such as in the data we use from the Thomas experiments, is difficult. Instead, we take the approach of specifying a plausible generative probabilistic world model, and we then construct the information partitions from this factored representation via a straightforward algorithm. The generative world model specifies the pieces of information, or "observations", that each player receives. The algorithm for generating information partitions, which is specified precisely in our supplementary materials, consists of iterating over all combinations of random values of variables in the generative world model, treating each such combination as a state in Ω, and for each player grouping together the states that yield identical observations.

Algorithm 1 common_p_belief(C, i, ω)

$E := \Omega$; $F := \Omega$
while $P_i(F \mid \omega) > 0$ **do**
  $p :=$ evidence_level(F, C)
  $E := F$
  $F :=$ super_p_evident(E, C, p)
$p :=$ evidence_level(E, C)
return p

Algorithm 2 evidence_level(E, C)

return $\min_{\omega \in E} \min$_belief(E, C, ω)

5.4.2 Computing Common p-Belief

Algorithms 1-4 present the functions needed to compute perceived maximal common p-belief. Formally, given a player i, a particular state ω, and an event C, Algorithm 1 computes the largest value p for which player i p-believes that there is common p-belief in C. Note that it is insufficient to compute the largest value of p for which there is common p-belief in C at ω, since in general at state ω player i only knows that the event $\Pi_i(\omega)$ has occurred, not that ω specifically has occurred. Relatedly, note that while Algorithm 1 takes ω as input for convenience, the algorithm only depends on $\Pi_i(\omega)$, and hence could be executed by a player.

Formal proofs of the correctness of these algorithms are included in our supplementary materials. The basic logic of the algorithms is to maintain a candidate p-evident event E, and to gradually remove elements from E to make it more p-evident until a point where player i believes the event to be impossible because the elements of the event are no longer consistent with that player's observations. By only removing elements that either cause E to be unlikely or cause C to be unlikely, we are guaranteed to arrive at a more p-evident event at each iteration, and one that preserves belief in C. By starting with E as the entire state space, the final candidate event must be the largest, most p-evident event that player i p-believes at state ω in which C is common p-belief. This algorithm can also be viewed as traversing a unique nested sequence of maximally evident events (independent of i and ω) induced by C and the structure of $(\Omega, \mu, (\Pi_0, \Pi_1))$, halting at the first event in this sequence that does not

Algorithm 3 min_belief(E, C, ω)

return $\min_{i \in \{0,1\}} \min(P_i(E \,|\, \omega), P_i(C \,|\, \omega))$

Algorithm 4 super_p_evident(E, C, p)

while E has changed **do**
  if $\exists\, \omega \in E : \text{min\_belief}(E, C, \omega) \leq p$ **then**
    $E := E \setminus \{\omega\}$
return E

include any elements of $\Pi_i(\omega)$.

The rational p-belief strategy consists of player i taking action A if common_p_belief($x = 1$, i, ω) $> \frac{c-b}{a-b}$. The matched p-belief strategy consists of player i choosing A with probability common_p_belief($x = 1$, i, ω).

5.4.3 Iterated Reasoning

We now present our algorithms for the iterated reasoning strategies. For level $k > 0$, given a player i and a state ω, the iterated maximization strategy computes

$$f_i^k(\omega) = \mathbb{1}\Big(\sum_{\omega' \in \Pi_i(\omega)} P_i(\omega' \,|\, \omega) \Big(P_i(x = 1 \,|\, \omega') f_{1-i}^{k-1}(\omega') a + P_i(x = 0 \,|\, \omega') f_{1-i}^{k-1}(\omega') d + \big(1 - f_{1-i}^{k-1}(\omega')\big) b \Big) > c \Big),$$

where $\mathbb{1}()$ is the indicator function, and $f_i^0(\omega) = \mathbb{1}(P_i(x = 1 \,|\, \omega) > \frac{c-b}{a-b})$. If $f_i^k(\omega) = 1$, then player i plays A, and otherwise player i plays B. For the iterated matching strategy,

$$q_i^k(\omega) = P_i(x = 1 \,|\, \omega) \cdot \sum_{\omega' \in \Pi_i(\omega)} P_i(\omega' \,|\, \omega) q_{1-i}^{k-1}(\omega').$$

A is then played with probability q_i^k, $q_i^0 = P_i(x = 1 \,|\, \omega)$.

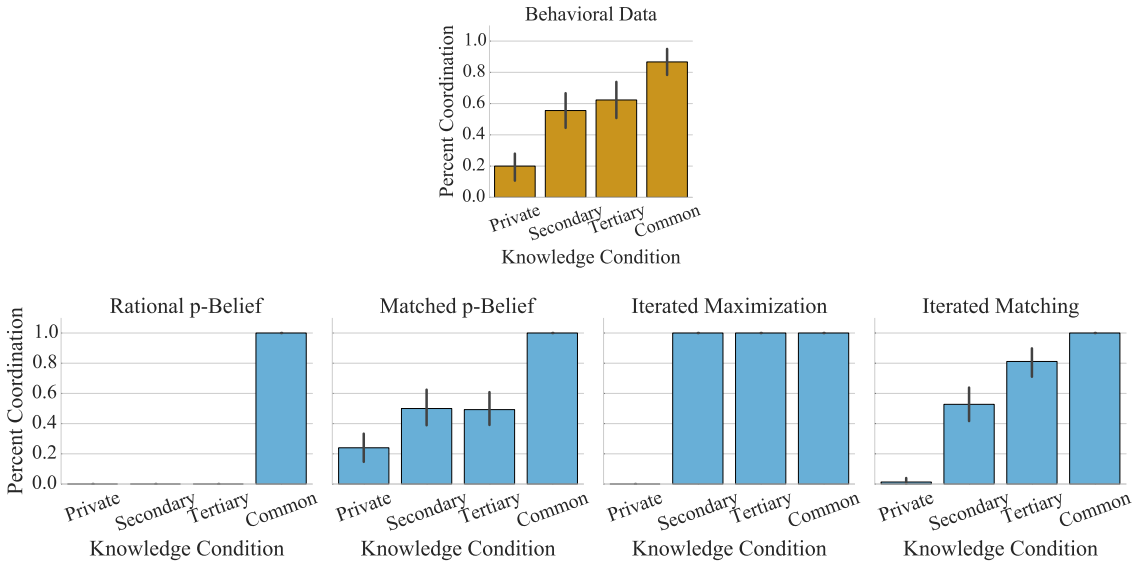

Figure 5-1: Data from the Thomas experiments and the predictions of each of the models we consider.

5.5 Data

We now present the data we use for our empirical results. The dataset comes from the Thomas experiments [113]. These experiments presented participants with a stylized coordinated attack problem couched in a story about the butcher and the baker of a town. In their story, these merchants can either work together to produce hot dogs, or they can work separately to produce chicken wings and dinner rolls, respectively. The merchants can sell chicken wings and dinner rolls separately for a constant profit of c each on any day, but the profit of hot dogs varies from day-to-day. The merchants make a profit of a each if $x = 1$ on a particular day or d if $x = 0$. There is also a loudspeaker that sometimes publicly announces the prices of hot dogs, and a messenger who runs around the town delivering messages. The experiments had four different knowledge conditions that specified the information that participants received:

1. **Private Knowledge:** "The Messenger Boy [sic] has not seen the Butcher today, so he cannot tell you anything about what the Butcher knows."

2. **Secondary Knowledge:** "The Messenger Boy says he stopped by the butcher

shop before coming to your bakery. He tells you that the Butcher knows what today's hot dog price is. However, he says that he forgot to mention to the Butcher that he was coming to see you, so the Butcher is not aware that you know today's hot dog price."

3. **Tertiary Knowledge:** "The Messenger Boy mentions that he is heading over to the butcher shop, and will let the Butcher know today's price as well. The Messenger Boy will also tell the Butcher that he just came from your bakery and told you the price. However, the Messenger Boy will not inform the Butcher that he told you he would be heading over there. So, while the Butcher is aware that you know today's price, he is not aware that you know that he knows that."

4. **Common Knowledge:** "The loudspeaker broadcast the market price . . . The messenger boy did not come by. Because the market price was broadcast on the loudspeaker, the Butcher knows today's price, and he knows that you know this information as well."

After being shown this information as well as additional information indicating that $x = 1$, the participants were asked whether they would like to try to make hot dogs or not. The dataset from this experiment is visualized in Figure 5-1. Since the researchers provided evidence that the behavior of participants in their two-player experiments was invariant to payoffs, here we focus on their first payoff condition, in which $a = 1.1$, $b = 0$, $c = 1$, and $d = 0.4$.

We use this dataset to test whether the coordination strategies we have described are good models of human coordination in this setting. In order to be able to generate predictions for these models, we must determine a state space that represents the story in the Thomas experiments. We designed the following two probabilistic generative world models (one for the messenger, and one for the loudspeaker) to be consistent with a reading of the knowledge conditions from those experiments. The observe(i,o)

function indicates that player i observes o.

Messenger:

$x \sim \text{Bernoulli}(\delta)$

$\text{visit}_0 \sim \text{Bernoulli}(0.5)$

$\text{visit}_1 \sim \text{Bernoulli}(0.5)$

$\text{tell\_plan}_0 \sim \text{visit}_0 \wedge \text{Bernoulli}(0.5)$

$\text{tell\_plan}_1 \sim \text{visit}_1 \wedge \text{Bernoulli}(0.5)$

if visit_0:

$\text{observe}(0, x)$

if tell_plan_0:

$\text{observe}(0, (\text{visit}_1, \text{tell\_plan}_1))$

if visit_1:

$\text{observe}(1, (x, \text{visit}_0))$

if tell_plan_1:

$\text{observe}(1, \text{tell\_plan}_0)$

Loudspeaker:

$x \sim \text{Bernoulli}(\delta)$

$\text{broadcast} \sim \text{Bernoulli}(0.5)$

if broadcast:

$\text{observe}(0, x), \text{observe}(1, x)$

These models share a free parameter δ. We take $\delta = 0.25$. This setting provides a closer fit to the empirical data than the maximum entropy setting of $\delta = 0.5$. We interpret statements that one player is "not aware" as meaning that the player could have been made aware, and assign a maximum entropy probability of 0.5 to these events.

The state spaces corresponding to these world models consist of the sets of all

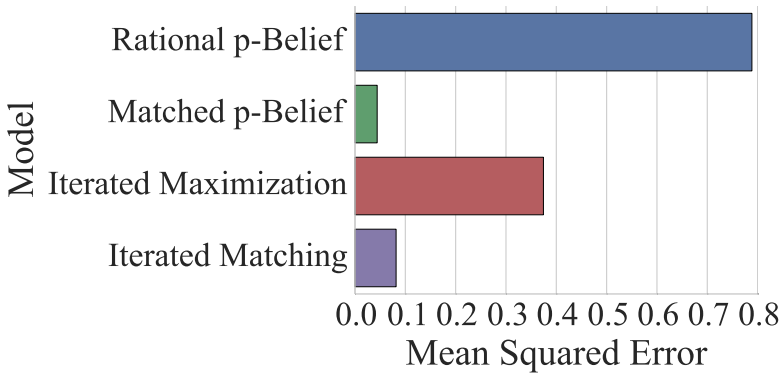

Figure 5-2: The mean-squared error of each model's predictions on the Thomas experiments' data.

possible combinations of variables in the models' generative processes: (x, visit$_0$, visit$_1$, tell_plan$_0$, tell_plan$_0$) for $\Omega_{\text{messenger}}$ and (x, broadcast) for $\Omega_{\text{loudspeaker}}$. The generative processes also uniquely specify probability measures over each state space. The knowledge conditions correspond to the following states. Private: $(1, 1, 0, 1, 0) \in \Omega_{\text{messenger}}$, Secondary: $(1, 1, 1, 0, 1) \in \Omega_{\text{messenger}}$, Tertiary: $(1, 1, 1, 1, 0) \in \Omega_{\text{messenger}}$, and Common Knowledge: $(1, 1) \in \Omega_{\text{loudspeaker}}$. The participants act as player 0 in all but the secondary condition. Due to high ambiguity in the wording of the private knowledge condition, we considered two plausible readings. Either the messenger is communicating an intention to not visit the other player, or the messenger is being unhelpful in not offering any information about the messenger's plan. By using the state $(1, 1, 0, 1, 0)$ we assume the first interpretation. This interpretation results in a better empirical fit.

5.6 Results

We now present our empirical results. We first examine the predictions of each of the coordination strategies we consider given the generative processes representing the Thomas experiments. We then examine the extent to which a computer agent equipped with the best fitting model of human coordination is able to achieve higher payoffs in a simulated human-agent coordination problem. All of our code is available

online at https://github.com/pkrafft/modeling-human-ad-hoc-coordination.

5.6.1 Model Comparison

To perform model comparison we compute the probability of choosing A that each model predicts given our formal representations of each of the four knowledge conditions. We then compare these predicted probabilities to the actual probabilities observed in the Thomas experiments. For the two iterated reasoning models we use a grid search over $[0, 1, 2, 3, 4, 5]$ to find the best fitting k for each model (ultimately $k = 1$ in iterated maximization and $k = 3$ in iterated matching). The specific predictions of each model are shown in Figure 5-1. As shown in Figure 5-2, the matched p-belief model achieves the lowest mean-squared error. Qualitatively, the most striking aspect of the data that the matched p-belief model successfully captures is the similarity in the probability of coordination between the secondary and tertiary knowledge conditions. The two models that involve maximizing agents (rational p-belief and iterated maximization) both make predictions that are too extreme. The iterated matching model offers a competitive second place fit to the data, but it fails to capture the similarity between the middle two knowledge conditions.

The reason that the matched p-belief model makes good predictions for the two middle conditions is that the player in both of those conditions has the same amount of uncertainty appearing at some level of that player's infinite recursive hierarchy of interpersonal beliefs. Common p-belief essentially represents a minimum taken over all these levels, and thus the common p-belief in each of those two conditions is the same. The rational p-belief model is aware of the uncertainty at higher levels of recursive belief, but its predictions are too coarse due to the assumption of utility maximization. An interesting avenue for future work is to examine whether the rational p-belief model can be relaxed in any other way to allow for intermediate predictions, such as by allowing for heterogeneity in interpretations of the world model across agents. It is possible that the matched p-belief model is approximating such a case.

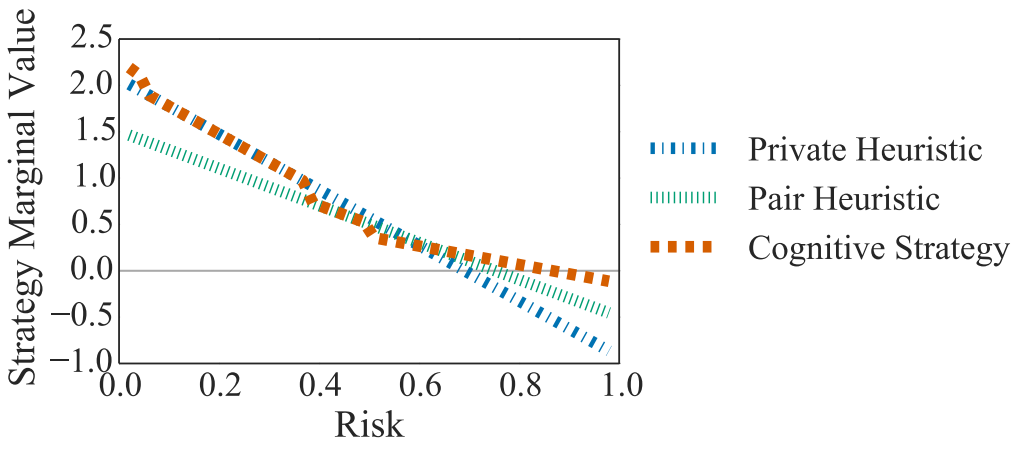

Figure 5-3: Performance of agents in our simulated human-agent coordination experiments. A strategy's marginal value is the expected sum of payoffs the strategy obtained in each of the four knowledge conditions, minus the payoffs that could have been obtained by always playing B.

5.6.2 Human-Agent Coordination

Besides testing the fit of the models of human coordination that we have proposed, we are also interested in whether the best fitting model helps us improve outcomes in human-agent coordination. We use the data from the Thomas experiments to evaluate this possibility. For this task our computer agents implement what we call a "cognitive strategy". An agent using the cognitive strategy chooses the action that maximizes expected utility under an assumption that the agent's companion is using the matched p-belief model. These agents play the humans' companion player in each of the four knowledge conditions of the Thomas experiments (player 1 in all but the Secondary condition). We evaluate the payoffs from the agents' actions using the human data from the Thomas experiments. In this simulation we vary the payoffs (a, b, c, d), and we assume that the humans would remain payoff invariant across the range of payoffs that we use. This assumption is reasonable given that participants in the Thomas experiments displayed payoff invariance in the two-agent case. We vary the payoffs according to risk, taking the payoffs to be $(1, 0, p^*, 0)$ for a particular risk level $p^* = \frac{c-b}{a-b}$. As shown in Figure 5-3, we find that having the matched p-belief model of human coordination may help in human-agent coordination. We compared to two baseline strategies: an agent using a "private heuristic" who always coordinates if the

agent knows $x = 1$, and an agent using a "pair heuristic" who always coordinates if the agent knows that both the agent and the human know $x = 1$. The private heuristic achieves good performance for low risk levels, and the pair heuristic achieves good performance for high risk levels. The cognitive strategy achieves good performance at both low and high levels of risk, and only has negative marginal value over always playing the safe action B at very high levels of risk.

5.7 Discussion

In the present chapter I focused on laying the groundwork for using common p-belief in AI and cognitive modeling. In particular, I developed an efficient algorithm for the inference of maximal perceived common p-belief, I showed that the coordination strategy of probability matching on common p-belief explains certain surprising qualitative features of existing data from a previous human experiment, and I showed that this model may also help improve agent outcomes in human-agent coordination. This work has three main limitations. Due to the small amount of human data I had and the lack of a held-out test set, my empirical results are necessarily only suggestive. While the data are inconsistent with the rational p-belief model and the iterated maximization model, the predictions of the iterated matching model and the matched p-belief model are both reasonably good. The strongest evidence we have favoring the matched p-belief model is this model's ability to produce equal amounts of coordination in the secondary and tertiary knowledge conditions as well as a low amount with private knowledge and a high amount with common knowledge. No iterated reasoning model under any formulation we could find of the Thomas experiments was able to capture the equality between the two middle conditions while maintaining good predictions at the extremes. Two other important limitations of my work are that the coordination task I consider did not involve intentional communication, and that the state and action spaces of the task were simple. While these features allowed me to easily test the predictions of each of my alternative models, it would be interesting to see how the models I considered would compare in more complex environments. A

related interesting direction for future work is the application of inference of common p-belief through reasoning about p-evident events to artificial distributed systems, such as for developing or analyzing bitcoin/blockchain-like protocols, synchronizing remote servers, or distributed planning in ad hoc multi-robot teams.

5.8 Proofs

These supplementary sections present the proofs of the correctness of our algorithms. Our key proofs are self-contained, but they rely on some preliminary results involving the structure of p-evident events in finite probability spaces. We first offer proofs of these preliminary results in the "Initial Results" section. We then prove that the rational p-belief strategy is an equilibrium in the "Models" section. Finally, we show the correctness of our algorithm for computing perceived maximal common p-belief in the "Algorithms" section. This final section also includes our algorithm for computing information partitions from a given probabilistic generative world model.

All of the proofs are original. Some of the less trivial lemmas and propositions may be novel contributions to the literature on p-evident events. In particular, we know of no work that explores the structure of what we will call "maximally evident C-indicating events", i.e. maximally p-evident events in which C is common knowledge at a given state ω. Our main result along these lines will be that any finite state space has a unique representation as a nested sequence of maximally evident C-indicating events. This result is central to understanding our algorithm for computing common p-belief.

5.8.1 Definitions

We first restate the most relevant definitions from our main text in a clearer format, and we give several additional definitions that will be helpful in our proofs. As described in the main text, we assume a Bayesian game with a finite state space $\Omega = \{\omega_1, \ldots, \omega_{|\Omega|}\}$ and information partitions Π_0 and Π_1, which are simply partitions of the state space Ω. The notation $\Pi_i(\omega)$ indicates the unique element of Π_i that

includes the state ω. These information partitions and a measure μ defined over Ω induce the conditional distribution $P_i(E \,|\, \omega) = \mu(E \,|\, \Pi_i(\omega)) = \frac{\sum_{\omega' \in E \cap \Pi_i(\omega)} \mu(\omega')}{\sum_{\omega' \in \Pi_i(\omega)} \mu(\omega')}$, which specifies the belief about an event E that the player i holds at state ω.

Definition 1 ([87]). *Player i is said to p-**believe** an event $E \subseteq \Omega$ at ω if $P_i(E \,|\, \omega) \geq p$.*

Definition 2 ([87]). *An event E is p-**evident** if for each $\omega \in E$, all players p-believe E at ω.*

Definition 3. *An event E is **super-p-evident** if for all $\omega \in E$ and for all players i, $P_i(E \,|\, \omega) > p$.*

Definition 4. *An event E is said to be a p-**evident C-indicating event** if E is a p-evident event and if $P_i(C \,|\, \omega) \geq p$ for all i and for all $\omega \in E$.*

Definition 5 ([87]). *The players have **common p-belief** in an event C at state ω if there exists a p-evident C-indicating event E that includes ω.*

Definition 6. *An event E is the **largest p-evident C-indicating event** (for a particular p) if E is a p-evident C-indicating event and for all events F, either F is not a p-evident C-indicating event, or $F \subseteq E$.*

Definition 7. *The **C-evidence level** of an event E is the maximum value of p for which E is a p-evident C-indicating event.*

Definition 8. *An event E is the **maximally evident C-indicating superset** of F if there exists a p such that p is the C-evidence level of E, E is the largest p-evident C-indicating event, and for all other events G, either G does not contain F or G has a lower C-evidence level than E.*

Definition 9. *An event E is the **largest super-evident C-indicating subset** of F if there exists a p such that p is the C-evidence level of F, E is a super-p-evident C-indicating event, and for all other events G, either G is not a subset of F, G is a subset of E, or the C-evidence level of G is less than or equal to p.*

5.8.2 Initial Results

Our first result is a simple lemma showing that the family of p-evident C-indicating events is closed under unions.

Lemma 1. *For a given event C, if E and F are p-evident C-indicating events, then $E \cup F$ is also a p-evident C-indicating event.*

Proof. Consider $\omega \in E \cup F$. For any player i, $P_i(E \cup F \,|\, \omega) \geq \min(P_i(E \,|\, \omega), P_i(F \,|\, \omega)) \geq p$ (since E and F are p-evident). Therefore $E \cup F$ is p-evident. Now suppose $\omega \in E$. In this case $P_i(C \,|\, \omega) \geq p$ since E is C-indicating. But also, if $\omega \in F$, then $P_i(C \,|\, \omega) \geq p$ since F is C-indicating. Therefore $E \cup F$ is C-indicating. □

Our next result is an immediate corollary and establishes the existence and uniqueness of largest p-evident C-indicating events for any p.

Corollary 1. *For a given event C, if there exists a p-evident C-indicating event E, then there exists a unique largest p-evident C-indicating event.*

Proof. Assume there exists a p-evident C-indicating event. Take $\mathcal{F}$ to be the set of all p-evident C-indicating events. By lemma 1, the set $G = \cup_{F \in \mathcal{F}} F$ consisting of the union of all of these events is also a p-evident C-indicating event. Therefore G is a p-evident C-indicating event containing all other p-evident C-indicating events. Moreover, E must be unique since it contains all other p-evident events that indicate C. □

The next lemma establishes a containment relationship between largest p-evident C-indicating events associated with different values of p.

Lemma 2. *For a given event C, if E is the largest p-evident C-indicating event, and F is the largest p'-evident C-indicating event, with $p \geq p'$, then $E \subseteq F$.*

Proof. Suppose E is the largest p-evident C-indicating event. Since $p' \leq p$, E is also a p'-evident C-indicating event. Thus E must be contained by the largest p'-evident C-indicating event. □

We now show that a unique maximally evident C-indicating event exists around any subset of Ω.

Lemma 3. *For a given event C, and for any event E, there exists a unique maximally evident C-indicating superset of E.*

Proof. Consider an event E. The C-evidence level of E is given by $p = \min_{i \in [0,1], \omega \in E} \min(P_i(E \mid \omega), P_i(C \mid$ and therefore exists. Moreover, since E is a p-evident C-indicating event, E must be contained by the unique largest p-evident C-indicating event F (which is guaranteed to exist by corollary 1). Further, by definition F must also contain all other p-evident supersets of E, and therefore F is the maximally evident C-indicating superset of E. □

The following lemma, which will be useful in proving our main theoretical result, shows that the family of maximally evident C-indicating events is highly constrained.

Lemma 4. *For a given event C, for any event E, there exists some $\omega \in E$ such that the maximally evident C-indicating superset of E is equal to the maximally evident C-indicating superset of $\{\omega\}$.*

Proof. Let F be the maximally evident C-indicating superset of E, guaranteed to exist by lemma 3. Let p_F be the C-evidence level of F. Consider an $\omega \in E$. Let G_ω be the maximally evident C-indicating superset of $\{\omega\}$, and let p_{G_ω} be the C-evidence level of G_ω. We must have that $p_{G_\omega} \geq p_F$ for all ω (Since F contains ω, $p_F > p_{G_\omega}$ would violate the fact that G_ω is maximally evident.) Suppose $p_{G_\omega} > p_F$ for all ω. Then, by lemma 1, $H = \cup_{\omega \in E} G_\omega$ would be a p_H-evident C-indicating event with $p_H \geq \min_\omega p_{G_\omega} > p_F$. However, H clearly contains E, and thus $p_H > p_F$ violates the fact that F is the maximally evident C-indicating superset of E. Therefore, there must exist some ω such that $p_{G_\omega} = p_F$. Since largest p-evident C-indicating events are unique by corollary 1, we must have $H = G_\omega$. □

The following theorem is our main theoretical result, and drives the correctness of our algorithms. This theorem, which is also illustrated in Figure 5-4, states that any

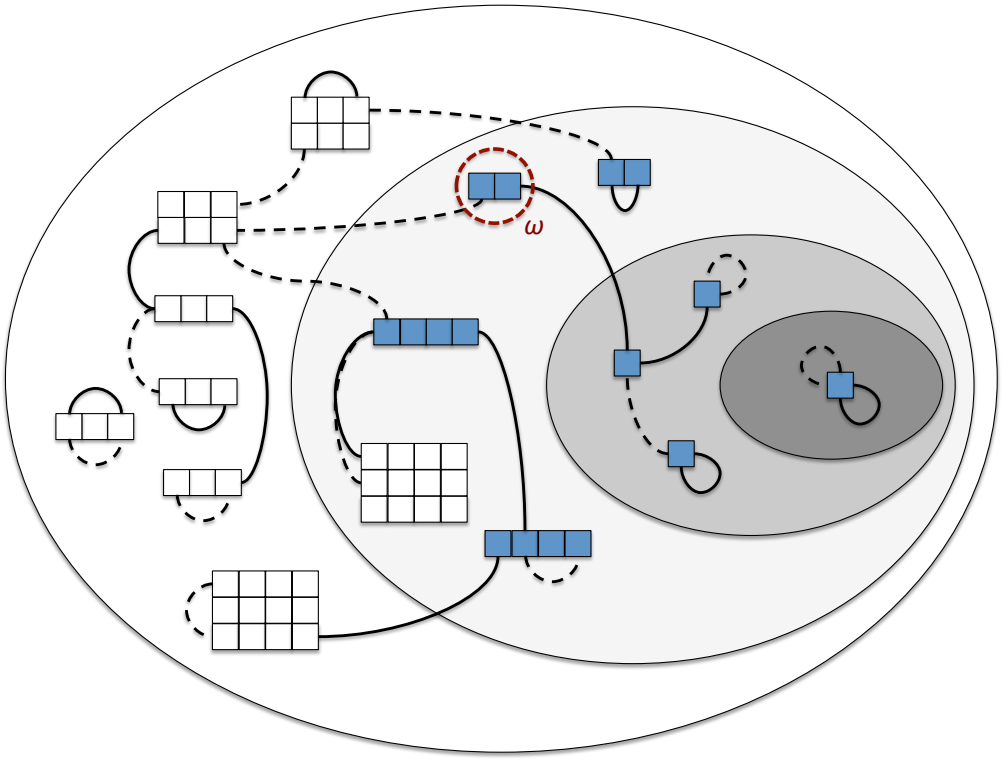

Figure 5-4: Any finite state space can be uniquely represented as a nested sequence of maximally evident C-indicating events. The diagram in this figure represents the generative process of the messenger described in the main text of this chapter (with $\delta = 0.25$). Each contiguous rectangle of blocks represents a state in Ω, and the measure of the state is given by the area of the rectangle. States that are shaded are members of $C = \{\omega \in \Omega : x(\omega) = 1\}$. The solid lines between states represent the information partition of player 0, while the the dotted lines represent that of player 1. Two states belong to the same element of a player's information partition if they are connected by some path in the graph induced by that player's edges. Self-loops indicate that a player has no uncertainty about the state when the state obtains. The four nested ovals are the four maximally evident C-indicating events in this state space ($\mathcal{F}$ in theorem 1), and the grey shading in the ovals represents the C-evidence levels of those events (0, 0.25, 0.5, and 1.0). Our algorithm iterates over these maximally evident events rather than over all possible events, and at ω for player i returns the C-evidence level associated with the last such event containing some element of $\Pi_i(\omega)$. For instance, at the circled state the algorithm will find the third nested event for player 0 and the second for player 1.

finite Ω has a unique representation as a nested sequence of maximally evident C-indicating events. The efficiency of our algorithm stems from only searching through this sequence of subsets, rather than all possible subsets, in order to compute common p-belief at any state ω.

Theorem 1. *Given an information structure $(\Omega, \mu, (\Pi_0, \Pi_1))$ with $|\Omega| < \infty$ and any event $C \subseteq \Omega$, let $\mathcal{F} = \{E'_1, \ldots, E'_n\}$ be the set of maximally evident C-indicating events of Ω, i.e. the set of events E'_i for which there exists some $F \subseteq \Omega$ such that E'_i is the maximally evident C-indicating superset of F. $\mathcal{F}$ can be ordered into a nested sequence of subsets, $E_1 \supset E_2 \supset \ldots \supset E_n$ such that E_i is the largest super-evident C-indicating subset event of E_{i-1} for any $i > 1$.*

Proof. We prove this theorem by construction. We know from lemma 3 that for each $\omega_j \in \Omega$, there exists a maximally evident C-indicating superset of $\{\omega_j\}$. We label this event E'_{ω_j}. Lemma 4 implies that collection of events $\{E'_{\omega_j} : \omega_j \in \Omega\}$ is equal to $\mathcal{F}$, the entire set of maximally evident C-indicating events. We also know from lemma 3 that each E'_{ω_j} is associated with a particular C-evidence level, which we will label p_{ω_j}. Without loss of generality, we can assume that the index j sorts these C-evidence levels by their magnitudes. This produces a finite non-decreasing sequence of values in $[0, 1]$: $p_{\omega_1} \leq p_{\omega_2} \leq \ldots \leq p_{\omega_{|\Omega|}}$. By lemma 2 the events E'_{ω_j} thus form a nested sequence of subsets: $E'_{\omega_1} \supseteq E'_{\omega_2} \supseteq \ldots \supseteq E'_{\omega_{|\Omega|}}$. We can collapse the events that are equal to each other to arrive at a sequence of strict subsets: $E_1 \supset E_2 \supset \ldots \supset E_n$.

Now consider E_i for some $i > 1$. By construction E_i is a super-$p_{\omega_{i-1}}$-evident C-indicating event, and $E_i \subset E_{i-1}$. Now take F to be another super-$p_{\omega_{i-1}}$-evident C-indicating event, and let p_F be the C-evidence level of F. If $p_F \geq p_{\omega_i}$, then we must have $E_i \supseteq F$ by lemma 2 since E_i is the largest p_{ω_i}-evident event. Now suppose $p_F < p_{\omega_i}$, so $p_{\omega_{i-1}} < p_F < p_{\omega_i}$. Let G be the maximally evident C-indicating superset of F, and let p_G be the C-evidence level of G. Clearly $p_G \geq p_F$, and by construction we must also have $p_G = p_{\omega_j}$ since $\mathcal{F}$ contains all maximally evident C-indicating supersets, and hence $G = E_j$, for some j. But since $p_{\omega_j} = p_G \geq p_F > p_{\omega_{i-1}}$, we must have $j \geq i$, and so $F \subseteq G = E_j \subseteq E_i$. Therefore again $F \subseteq E_i$. Hence E_i must be

the largest C-indicating subset of E_{i-1}. □

The following corollary to this theorem provides the basis for our iterative algorithm.

Corollary 2. *Given a maximally evident C-indicating event E, either there exists a unique largest super-evident C-indicating subset of E, or no super-p-evident C-indicating subsets exist.*

Proof. Since E is a maximally evident C-indicating event, E is equal to some E_i in the sequence $\mathcal{F}$ given by theorem 1. Therefore, if $i < n$, E_{i+1} is the unique largest super-evident C-indicating subset of E or no such subsets exist, as indicated in theorem 1. □

Finally, we have four technical lemmas that will be useful for analyzing our specific algorithms and player strategies. The first lemma is a simple constraint on belief about p-evident events that follows from how information partitions work.

Lemma 5. *For any p-evident event E, either $P_i(E \,|\, \omega) = 0$ or $P_i(E \,|\, \omega) \geq p$ for all ω and for all i.*

Proof. If $P_i(E \,|\, \omega) > 0$, then there must exist an $\omega' \in \Pi(\omega)$ such that $\omega' \in E$. Since E is p-evident, then $P_i(E \,|\, \omega') \geq p$. But then since $\Pi_i(\omega) = \Pi_i(\omega')$ by the definition of an information partition, $P_i(E \,|\, \omega) = \mu(E \,|\, \Pi_i(\omega)) = \mu(E \,|\, \Pi_i(\omega')) = P_i(E \,|\, \omega')$, and hence $P_i(E \,|\, \omega) \geq p$. □

The second lemma connects the containment relationships between p-evident events to the relationship between beliefs about those events.

Lemma 6. *If player i p-believes a p-evident C-indicating event, then player i p-believes the largest p-evident C-indicating event.*

Proof. Let F be a p-evident C-indicating event and let E be the largest p-evident C-indicating event. Since E is largest, $F \subseteq E$, so we must have $P_i(E \,|\, \omega) = P_i(E \cup F \,|\, \omega) \geq P_i(F \,|\, \omega)$. Hence if i p-believes F, i must also p-believe E. □

The next lemma states that p-belief is transitive.

Lemma 7. *If player i p-believes F at all elements of an event E, and player i p-believes G at all elements of F, then player i p-believes G at all elements of E.*

Proof. Let ω be an element of E. Since i p-believes F at ω, $P_i(F \mid w) \geq p$. Therefore we must have some $\omega' \in \Pi_i(\omega) \cap F$. Since player i p-believes G at all states in F, we must have $P_i(G \mid \omega') \geq p$. By the definition of an information partition, we must then also have $P_i(G \mid \omega) = P_i(G \mid \omega') \geq p$. Hence player i p-believes G at all $\omega \in E$. □

The final lemma states, roughly, that with regard to p-evident events a player's beliefs about another player cannot be too inconsistent with the first player's own beliefs.

Lemma 8. *If player i p-believes that player $1-i$ p-believes the largest p-evident C-indicating event E, and player i p-believes C, then player i must p-believe E.*

Proof. Consider an arbitrary $\omega \in \Omega$ and an arbitrary p. Take the largest p-evident C-indicating event E. Suppose at ω player i p-believes that player $1-i$ p-believes E, and player i p-believes C. To generate the first belief, there must be an element of the information partition of player $1-i$, $T \in \Pi_{1-i}$, such that $P_{1-i}(E \mid T) \geq p$, and another set of states $S \subseteq \Pi_i(\omega) \cap T$ such that $P_i(S \mid \omega) \geq p$. But then, since states within a player's information partition are indistinguishable and since $S \subseteq \Pi_i(\omega) \cap T$, we must have for all $\omega' \in S$, $P_{1-i}(E \mid \omega') = P_{1-i}(E \mid T) \geq p$ and $P_i(S \mid \omega') = P_i(S \mid \omega) \geq p$. Therefore, since E is p-evident, $S \cup E$ is also a p-evident event (if $\omega'' \in S$, then player $1-i$ p-believes E and player i p-believes S, while if $\omega'' \in E$, then both players p-believe E). Moreover, since player i p-believes C at ω (and therefore at $\Pi(\omega)$ and S) and since $1-i$ p-believes C at E (and hence at S by lemma 7), then both players p-believe C at all states in $S \cup E$. Therefore $S \cup E$ is a p-evident C-indicating event (and, moreover, player i p-believes $S \cup E$ at ω since player i p-believes S at ω). But E was assumed to be the largest p-evident C-indicating event, so we must have $S \cup E = E$. Hence player i p-believes E. □

5.8.3 Models

Strategic Coordination

In this section we show the rational p-belief strategy forms an equilibrium in the coordination game we study. Recall the rational p-belief strategy is that player i plays action A if and only if player i believes with at least probability $p^* = \frac{c-b}{a-b}$ that both players have common p^*-belief that $x = 1$.

Proposition 1. *Assuming noiseless messages and assuming $p^* > \mu(x = 1)$, the rational p-belief strategy maximizes the expected return of player i at every $\omega \in \Omega$, given player $1 - i$ also uses the same strategy.*

Proof. Suppose that messages are noiseless, i.e. $P_i(x = 1 \,|\, \omega) > P_i(x = 1) \Rightarrow P_i(x = 1 \,|\, \omega) = 1$ for all ω. Take $\omega \in \Omega$.

Take F to be some p^*-evident $(x = 1)$-indicating event. Take E to be the largest such event. Suppose player i p^*-believes F. By lemma 6, player i also p^*-believes E. By the definition of p-evident events, for any $\omega' \in E$, we must then have that player $1 - i$ also p^*-believes E at ω'. Therefore by lemma 7 player i also p^*-believes that player $1-i$ p^*-believes E. Since player $1-i$ is assumed to be using the rational p-belief strategy, player i therefore p^*-believes player $1-i$ will play A. Since we have assumed $p^* > P_i(x = 1)$ and messages are noiseless, we must have $P_i(x = 1 \,|\, \Pi_i(\omega)) = 1$. Then the expected return for i of playing A must be at least $p^* \cdot a + (1-p^*)b = \frac{c-b}{a-b}a + \frac{a-c}{a-b}b = c$. Therefore, playing A maximizes player i's expected return (and if player i p'-believes a p'-evident $(x = 1)$-indicating event, $p' > p^*$, then the expected return from A is strictly greater than for B).

Now suppose player i p^*-believes $x = 1$ but that there is no p^*-evident $(x = 1)$-indicating event F such that $P_i(F \,|\, \omega) \geq p^*$. Take E to be the largest p^*-evident $(x = 1)$-indicating event. Then player i cannot p^*-believe that player $1 - i$ p^*-believes E (since otherwise player i would p^*-believe E by lemma 8). Thus, since player $1 - i$ is assumed to be using the rational p-belief strategy, player i believes with probability at most $p' < p^*$ that player $1-i$ will play A. Then since $p' \cdot a + (1-p')b <$

$p^* \cdot a + (1 - p^*)b = c$, the expected utility from playing A to player i is less than from playing B. (And if i does not p^*-believe $x = 1$, the same arithmetic holds.)

□

5.8.4 Algorithms

Computing Information Partitions

In this section we state the simple algorithm for converting the generative process descriptions of the probabilistic models in our main text to information partitions. The "run(i, ω)" function takes as input a player i and a state ω expressed as a tuple, uses the variables in the tuple for each random draw within the generative process, and returns a composition of the *observe* calls for player i (For example, run(0, (1,1,0,1,0)) would return [(1),(0,0)] in the messenger model). Our algorithm groups together states with the same observations into elements of the information partitions.

Algorithm 5 information_partition(i)

```
partition := dict()
for ω ∈ Ω do
    obs := run(i, ω)
    if obs ∈ partition then
        partitions[obs].append(ω)
    else
        partitions[obs] := [ω]
return partitions
```

Computing Common p-Belief

We now prove the correctness of the common_p_belief algorithm given in the main text. We first prove a lemma stating the correctness of the super_p_evident algorithm.

Lemma 9. *Assuming E is a maximally evident C-indicating event, and assuming p is the C-evidence level of E, then the set returned by evaluating "super_p_evident(E, p)"*

is the largest super-evident C-indicating subset of E, or the empty set if and only if such an event does not exist.

Proof. First note that from the condition in the "if" statement of the function, all states that remain in E when the function returns will have the properties $P_i(E \,|\, \omega) > p$ and $P_i(C \,|\, \omega) > p$. Therefore, the function only returns super-p-evident C-indicating subsets of E (and since p is the C-evidence level of E, strict subsets must be returned), or the empty set if no super-p-evident subsets exists. Next, note that the "while" loop only removes elements of E that cannot belong to the largest super-evident C-indicating subset of E, if such an event exists. We can see this by induction on the items removed. Suppose that the first ω removed belonged to F, the largest super-evident C-indicating subset of E. Then we would have $P_i(F \,|\, \omega) > p$ since F must be super-p-evident. But since $F \subseteq E$, by lemma 2 we would then also have $P_i(E \,|\, \omega) > p$, which contradicts the fact that ω was removed. Now assume that the first k elements removed do not belong to F, and let E' be the set E minus those elements. Then suppose the $(k+1)$st element removed, if it exists, belonged to F. Since all the previous elements removed did not belong to F by the inductive assumption, $F \subseteq E'$. But then analogous to the base case, the $(k+1)$st element, ω_{k+1}, belonging to F implies $P_i(E' \,|\, \omega_{k+1}) > p$, which contradicts the fact that this element was removed. Therefore this function returns exactly the largest super-evident C-indicating subset of E. Finally, since Ω is assumed to be finite, and at least one element of E is removed in each iteration of the while loop, the function must terminate. □

Lastly, we show the correctness of our main algorithm.

Proposition 2. *The value returned by evaluating "common_p_belief(C, i, ω)" is the C-evidence level of the maximally evident C-indicating event E containing some element of $\Pi_i(\omega)$.*

Proof. First note "evidence_level(F, C)" computes the C-evidence level of F (which must exist by lemma 3). Also note Ω is the maximally evident C-indicating superset of Ω. Thus by lemma 9, the calls to "super_p_evident()" iteratively return the nested sequence of subsets of maximally evident subsets described in theorem 1.

If there does not exist a p-evident C-indicating event for $p > 0$, then by lemma 9 the first call to super_p_evident will return the empty set. In this case the C-evidence level of Ω will be 0, and hence the function will return 0.

Since the "while" loop will continue iterating until either F is empty or F is an event that player i believes to be impossible (i.e., that does not contain any elements in $\Pi_i(\omega)$), the last E before either of these cases occurred must be the maximally evident C-indicating event containing some element of $\Pi_i(\omega)$. The call to evidence_level then computes the C-evidence level of E. □

Our final result interprets the last result in terms of common-p-belief.

Corollary 3. *The value returned by evaluating "common_p_belief(C, i, ω)" is the maximum value of p such that player i p-believes there is common-p-belief in C at ω.*

Proof. By proposition 2, common_p_belief(C, i, ω) returns the C-evidence level of some event E that player i believes with probability greater than 0. Let p equal the C-evidence level of E. By lemma 5, player i must therefore p-believe E. Player i thus p-believes there is common-p-belief in C at ω since E is a p-evident C-indicating event. Now suppose there was some $p' > p$ such that player i p'-believes that there is common-p'-belief in C. Then by definition of common-p-belief there must exist a p'-evident event F that player i p'-believes. But then player i must also p'-believe the maximally evident C-indicating superset of F. However, this contradicts the fact that p was returned by common_p_belief(C, i, ω), since if player i p'-believes F, then F is a maximally evident event containing some element of $\Pi_i(\omega)$. □

Chapter 6

Conclusion

In this thesis I investigated the question of how disparate individuals combine into coherent groups. I presented a theoretical answer to this question: that three conditions suffice for rational collective agency, and that people are endowed with abilities, mechanisms, and propensities, perhaps biologically or culturally evolved, that allow us to achieve these conditions. I then focused on understanding the mechanisms that people use to address one of the three conditions, the formation of accurate shared beliefs.

In investigating specific mechanisms that people use for shared belief formation, a major theme arose. Theory of mind appears to be critical to the solutions that people have discovered to address the fundamental computational challenges of establishing rational collective agency. This finding informs another puzzle in the literature on collective intelligence—theory of mind as measured by emotional intelligence has been reliably associated with collective intelligence in groups [126, 31]. This result is also related to hypotheses about the importance of theory of mind to the uniqueness of the human species among animals [115].

The overall picture that emerges from these investigations is almost mystical in its beauty. People are connected by nothing. We are isolated independent biological agents who act alone. We have no physical connections and sparse biological connections to each other. Our saving grace is our skill as mind readers. Our "psychic" powers allow us to spontaneously create highly robust rational collective agency out of

highly heterogeneous people. We infer each others preferences, beliefs, and intentions. We imagine mechanisms to solidify our connections in our minds.

Another theme that emerged was the prominence of probability matching—a widely observed behavioral decision-making strategy. In the real-world case study I presented, probability matching was critical to the emergence of collective intelligence. Acting stochastically communicates more information than acting deterministically based on local optimization, and thereby allows for efficient information aggregation at the collective level.

A final major theme was the use of distributed computation to understand collective behavior. Researchers recently began using distributed algorithms to understand natural systems, such as networks of cells and colonies of ants and bees [32]. My work can be viewed as extending this line of work further into the study of human collective behavior. Distributed computation provides us with an elegant solution to the micro-macro problem because macro-level computation can provably be shown to emerge from micro-level algorithms. I provided several examples of this line of reasoning. A specific class of distributed algorithms that seems to hold promise for understanding coherent group behavior and collective intelligence is algorithms for distributed Baysian inference. I provided two examples of heuristic mechanisms that people use implicitly implementing Baysian inference in a distributed capacity. In other recent work not described in this thesis, I have begun investigating how to extend this framework of collective intelligence as distributed Bayesian inference to more complex real-world examples, such as rumor spreading during collective sensemaking [64].

The examples that I studied in this thesis also make interesting predictions about collective intelligence in other sociotechnical systems. Other work I have completed that is not described in this thesis has validated these predictions experimentally in the laboratory and the field. In particular, the social sampling model I introduced predicts that naïve herding leads to fragility in the dynamics of popularity. Together with Nicolás Della Penna, I tested this prediction in a set of cryptocurrency markets using a new experimental methodological [68], showing the first experimental evi-

dence for herding in financial markets [66]. The social sampling model also predicts that if popularity is well-calibrated, then social influence can lead to high levels of collective intelligence. Together with L. Elisa Celis, I confirmed this prediction in laboratory versions of online educational courses on computer science and art [20]. Future work could attempt to apply the insights from these predictions for engineering new sociotechnical systems.

Bibliography

[1] Shipra Agrawal and Navin Goyal. Analysis of Thompson sampling for the multi-armed bandit problem. In *25th Annual Conference on Learning Theory*, volume 23, pages 39.1–39.26. JMLR: Workshop and Conference Proceedings, 2012.

[2] Floyd H Allport. The group fallacy in relation to social science. *American Journal of Sociology*, 29(6):688–706, 1924.

[3] Jean-Marc Amé, José Halloy, Colette Rivault, Claire Detrain, and Jean Louis Deneubourg. Collegial decision making based on social amplification leads to optimal group formation. *Proceedings of the National Academy of Sciences*, 103(15):5835–5840, 2006.

[4] Sinan Aral and Christos Nicolaides. Exercise contagion in a global social network. *Nature Communications*, 8, 2017.

[5] Manuel Arellano. Computing robust standard errors for within-groups estimators. *Oxford Bulletin of Economics and Statistics*, 49(4):431–434, 1987.

[6] Kenneth J Arrow. A difficulty in the concept of social welfare. *Journal of Political Economy*, 58(4):328–346, 1950.

[7] Solomon E Asch. Opinions and social pressure. In *Readings about the Social Animal*, volume 193, pages 17–26. Worth Publishers New York, NY, 1955.

[8] Robert J Aumann. Agreeing to disagree. *The Annals of Statistics*, pages 1236–1239, 1976.

[9] A. Berdahl, C. J. Torney, C. C. Ioannou, J. J. Faria, and I. D. Couzin. Emergent Sensing of Complex Environments by Mobile Animal Groups. *Science*, 339(6119), 2013.

[10] Peter Berger and Thomas Luckmann. *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. Anchor Books, 1966.

[11] José M Bernardo and Adrian FM Smith. Bayesian theory, 2001.

[12] Daniel S Bernstein, Shlomo Zilberstein, and Neil Immerman. The complexity of decentralized control of markov decision processes. In *Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence*, pages 32–37, 2000.

[13] Ginestra Bianconi and Albert-László Barabási. Bose-Einstein condensation in complex networks. *Physical Review Letters*, 86(24):5632, 2001.

[14] Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, 100(5):992–1026, 1992.

[15] Robert Boyd, Peter J Richerson, and Joseph Henrich. The cultural niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences*, 108:10918–10925, 2011.

[16] Michael E Bratman. Shared cooperative activity. *The Philosophical Review*, pages 327–341, 1992.

[17] Colin Camerer. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton University Press, 2003.

[18] Ciro Cattuto, Alain Barrat, Andrea Baldassarri, Gregory Schehr, and Vittorio Loreto. Collective dynamics of social annotation. *Proceedings of the National Academy of Sciences*, 106(26):10511–10515, 2009.

[19] Elisa Celis, Peter Krafft, and Nisheeth Vishnoi. A distributed learning dynamics in social groups. In *ACM Symposium on Principles of Distributed Computing (PODC)*, 2017.

[20] L Elisa Celis, Peter M Krafft, and Nathan Kobe. Sequential voting promotes collective discovery in social recommendation systems. In *Tenth International AAAI Conference on Web and Social Media*, 2016.

[21] Christophe Chamley. *Rational Herds: Economic Models of Social Learning*. Cambridge University Press, 2004.

[22] Olivier Chapelle and Lihong Li. An empirical evaluation of thompson sampling. In *Advances in Neural Information Processing Systems*, pages 2249–2257, 2011.

[23] Yiling Chen, Arpita Ghosh, Michael Kearns, Tim Roughgarden, and Jennifer Wortman Vaughan. Mathematical foundations for social computing. *Communications of the ACM*, 59(12):102–108, 2016.

[24] Judith A. Chevalier and Dina Mayzlin. The Effect of Word of Mouth on Sales: Online Book Reviews. *Journal of Marketing Research*, 43(3):345–354, 2006.

[25] Robert B Cialdini and Noah J Goldstein. Social influence: Compliance and conformity. *Annual Review of Psychology*, 55:591–621, 2004.

[26] Iain D Couzin. Collective minds. *Nature*, 445(7129):715–715, 2007.

[27] Iain D. Couzin. Collective cognition in animal groups. *Trends in Cognitive Sciences*, 13(1):36–43, 2009.

[28] Nuh Aygun Dalkiran, Moshe Hoffman, Ramamohan Paturi, Daniel Ricketts, and Andrea Vattani. Common knowledge and state-dependent equilibria. In *Algorithmic Game Theory*, pages 84–95. Springer, 2012.

[29] Morris H DeGroot. Reaching a consensus. *Journal of the American Statistical Association*, 69(345):118–121, 1974.

[30] Émile Durkheim. *The Rules of Sociological Method*. 1895.

[31] David Engel, Anita Williams Woolley, Lisa X. Jing, Christopher F. Chabris, and Thomas W. Malone. Reading the Mind in the Eyes or Reading between the Lines? Theory of Mind Predicts Collective Intelligence Equally Well Online and Face-To-Face. *PLoS ONE*, 9(12):e115212, 2014.

[32] Ofer Feinerman and Amos Korman. Theoretical distributed computing meets biology: A review. In *Distributed Computing and Internet Technology*, pages 1–18. Springer, 2013.

[33] Asaf Frieder, Raz Lin, and Sarit Kraus. Agent-human coordination with communication costs under uncertainty. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*, pages 1281–1282, 2012.

[34] Ya'akov Gal and Avi Pfeffer. Networks of influence diagrams: A formalism for representing agents' beliefs and decision-making processes. *Journal of Artificial Intelligence Research*, 33(1):109–147, 2008.

[35] Samuel J Gershman, Eric J Horvitz, and Joshua B Tenenbaum. Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, 349(6245):273–278, 2015.

[36] Herbert Gintis. *The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences*. Princeton University Press, 2009.

[37] Piotr J Gmytrasiewicz and Edmund H Durfee. Decision-theoretic recursive modeling and the coordinated attack problem. In *Proceedings of the First International Conference on Artificial Intelligence Planning Systems*, pages 88–95, 1992.

[38] Robert L. Goldstone and Todd M. Gureckis. Collective Behavior. *Topics in Cognitive Science*, 1(3):412–438, 2009.

[39] Robert L. Goldstone, Michael E. Roberts, Winter Mason, and Todd Gureckis. Collective search in concrete and abstract spaces. In *Decision Modeling and Behavior in Complex and Uncertain Environments*, pages 277–308. Springer, 2008.

[40] Benjamin Golub and Matthew O Jackson. Naive learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics*, pages 112–149, 2010.

[41] Thomas L Griffiths, Charles Kemp, and Joshua B Tenenbaum. Bayesian models of cognition. In *The Cambridge Handbook of Computational Psychology*. Cambridge University Press, 2008.

[42] Thomas L Griffiths, Falk Lieder, and Noah D Goodman. Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic. *Topics in Cognitive Science*, 7(2):217–229, 2015.

[43] Thomas L Griffiths and Joshua B Tenenbaum. Optimal predictions in everyday cognition. *Psychological Science*, 17(9):767–773, 2006.

[44] Barbara J Grosz and Sarit Kraus. Collaborative plans for complex group action. *Artificial Intelligence*, 86(2):269–357, 1996.

[45] Joseph Y Halpern and Yoram Moses. Knowledge and common knowledge in a distributed environment. *Journal of the ACM*, 37(3):549–587, 1990.

[46] Ward A. Hanson and Daniel S. Putler. Hits and misses: Herd behavior and online product popularity. *Marketing Letters*, 7(4):297–305, 1996.

[47] Robert X. D. Hawkins. Conducting real-time multiplayer experiments on the web. *Behavior Research Methods*, 2014.

[48] Peter Hedström and Peter Bearman. *The Oxford Handbook of Analytical Sociology*. Oxford University Press, 2009.

[49] Richard J Herrnstein. Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, 4(3):267, 1961.

[50] Eshcar Hillel, Zohar S Karnin, Tomer Koren, Ronny Lempel, and Oren Somekh. Distributed exploration in multi-armed bandits. In *Advances in Neural Information Processing Systems*, pages 854–862, 2013.

[51] Thomas T Hills, Peter M Todd, David Lazer, A David Redish, Iain D Couzin, Cognitive Search Research Group, et al. Exploration versus exploitation in space, mind, and society. *Trends in Cognitive Sciences*, 19(1):46–54, 2015.

[52] Mark K. Ho, James MacGlashan, Amy Greenwald, Michael L. Littman, Elizabeth M. Hilliard, Carl Trimbach, Stephen Brawner, Joshua B. Tenenbaum, Max Kleiman-Weiner, and Joseph L. Austerweil. Feature-based joint planning and norm learning in collaborative games. In *Proceedings of the 38th Annual Conference of the Cognitive Science Society*, 2016.

[53] James Hollan, Edwin Hutchins, and David Kirsh. Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction*, 7(2):174–196, 2000.

[54] John A. Howard and Jagdish N. Sheth. *The Theory of Buyer Behavior*. MIT Press, 1969.

[55] Edwin Hutchins. *Cognition in the Wild*. MIT Press, 1995.

[56] Nicholas R Jennings, Luc Moreau, D Nicholson, S Ramchurn, S Roberts, T Rodden, and Alex Rogers. Human-agent collectives. *Communications of the ACM*, 57(12):80–88, 2014.

[57] Demetri Kantarelis. *Theories of the Firm*. Geneve: Inderscience, 2007.

[58] Emilie Kaufmann, Nathaniel Korda, and Rémi Munos. Thompson sampling: An asymptotically optimal finite-time analysis. In *Algorithmic Learning Theory*, pages 199–213. Springer, 2012.

[59] Jaya Kawale, Hung H Bui, Branislav Kveton, Long Tran-Thanh, and Sanjay Chawla. Efficient thompson sampling for online matrix-factorization recommendation. In *Advances in Neural Information Processing Systems*, pages 1297–1305, 2015.

[60] Michael Kearns. Experiments in social computation. *Communications of the ACM*, 55(10):56–67, 2012.

[61] Michael Kearns, Siddharth Suri, and Nick Montfort. An experimental study of the coloring problem on human subject networks. *Science*, 313(5788):824–827, 2006.

[62] Norbert L. Kerr and R. Scott Tindale. Group Performance and Decision Making. *Annual Review of Psychology*, 55(1):623–655, 2004.

[63] Max Kleiman-Weiner, MK Ho, JL Austerweil, Michael L Littman, and Josh B Tenenbaum. Coordinate to cooperate or compete: abstract goals and joint intentions in social interaction. In *Proceedings of the 38th Annual Conference of the Cognitive Science Society*, 2016.

[64] Peter Krafft, Kaitlyn Zhou, Isabelle Edwards, Kate Starbird, and Emma S Spiro. Centralized, parallel, and distributed information processing during collective sensemaking. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 2976–2987. ACM, 2017.

[65] Peter M Krafft, Chris L Baker, Alex Sandy Pentland, and Joshua B Tenenbaum. Modeling human ad hoc coordination. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, pages 3740–3746. AAAI Press, 2016.

[66] Peter M Krafft, Nicolás Della Penna, and Alex Pentland. Social herding in an online marketplace: Evidence from 100,000 tiny random trades, Under review.

[67] Peter M Krafft, Robert XD Hawkins, Alex Pentland, Noah Goodman, and Joshua B Tenenbaum. Emergent collective sensing in human groups. *Annual Conference of the Cognitive Science Society (CogSci)*, 2015.

[68] Peter M Krafft, Michael Macy, and Alex Pentland. Bots as virtual confederates: Design and ethics. *The 20th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW)*, 2017.

[69] Peter M Krafft, Julia Zheng, Wei Pan, Nicolás Della Penna, Yaniv Altshuler, Erez Shmueli, Joshua B Tenenbaum, and Alex Pentland. Human collective intelligence as distributed bayesian inference. *arXiv preprint arXiv:1608.01987*, 2016.

[70] Coco Krumme, Manuel Cebrian, Galen Pickard, and Alex Pentland. Quantifying social influence in an online cultural market. *PLoS ONE*, 7(5):e33785, 2012.

[71] Christopher Krupenye, Fumihiro Kano, Satoshi Hirata, Josep Call, and Michael Tomasello. Great apes anticipate that other individuals will act according to false beliefs. *Science*, 354(6308):110–114, 2016.

[72] David Lazer and Allan Friedman. The network structure of exploration and exploitation. *Administrative Science Quarterly*, 52(4):667–694, 2007.

[73] Barbara Levitt and James G March. Organizational learning. *Annual Review of Sociology*, 14(1):319–338, 1988.

[74] Christian List and Philip Pettit. Aggregating sets of judgments: An impossibility result. *Economics & Philosophy*, 18(1):89–110, 2002.

[75] Lars Ljungqvist and Thomas J Sargent. *Recursive Macroeconomic Theory*. MIT Press, 2012.

[76] Nancy A Lynch. *Distributed Algorithms*. Morgan Kaufmann, 1996.

[77] Thomas W Malone and Kevin Crowston. What is coordination theory and how can it help design cooperative work systems? In *Proceedings of the 1990 ACM Conference on Computer-Supported Cooperative Work*, pages 357–370. ACM, 1990.

[78] James G March. Exploration and exploitation in organizational learning. *Organization Science*, 2(1):71–87, 1991.

[79] David Marr. *Vision: A Computational Approach*. MIT Press, 1982.

[80] Winter Mason and Duncan J Watts. Collaborative learning in networks. *Proceedings of the National Academy of Sciences*, 109(3):764–769, 2012.

[81] Julian John McAuley and Jure Leskovec. From amateurs to connoisseurs: Modeling the evolution of user expertise through online reviews. In *Proceedings of the 22nd International Conference on World Wide Web*, pages 897–908. ACM, 2013.

[82] R. McElreath, A. V Bell, C. Efferson, M. Lubell, P. J Richerson, and T. Waring. Beyond existence and aiming outside the laboratory: estimating frequency-dependent and pay-off-biased social learning strategies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1509):3515–3528, 2008.

[83] Miller McPherson, Lynn Smith-Lovin, and James M Cook. Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1):415–444, 2001.

[84] Jessica R Mesmer-Magnus and Leslie A DeChurch. Information sharing and team performance: a meta-analysis. *Journal of Applied Psychology*, 94(2):535, 2009.

[85] Stanley Milgram, Leonard Bickman, and Lawrence Berkowitz. Note on the drawing power of crowds of different size. *Journal of Personality and Social Psychology*, 13(2):79, 1969.

[86] Susan Mohammed, Lori Ferzandi, and Katherine Hamilton. Metaphor no more: A 15-year review of the team mental model construct. *Journal of Management*, 36(4):876–910, 2010.

[87] Dov Monderer and Dov Samet. Approximating common knowledge with common beliefs. *Games and Economic Behavior*, 1(2):170–190, 1989.

[88] Stephen Morris and Hyun Song Shin. Approximate common knowledge and co-ordination: Recent lessons from game theory. *Journal of Logic, Language and Information*, 6(2):171–190, 1997.

[89] L. Muchnik, S. Aral, and S. J. Taylor. Social Influence Bias: A Randomized Experiment. *Science*, 341(6146):647–651, 2013.

[90] Allen Newell. The knowledge level: Presidential address. *AI Magazine*, 2(2):1, 1981.

[91] Pedro A. Ortega. *A unified framework for resource-bounded autonomous agents interacting with unknown environments*. PhD thesis, University of Cambridge, 2011.

[92] Wei Pan, Yaniv Altshuler, and Alex Pentland. Decoding social influence and the wisdom of the crowd in financial trading network. In *International Conference on Social Computing*, pages 203–209, 2012.

[93] John W Payne. Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance*, 16(2):366–387, 1976.

[94] Alex Pentland. *Social Physics: How Good Ideas Spread-The Lessons from a New Science*. Penguin, 2014.

[95] Stephen C. Pratt and David JT Sumpter. A tunable algorithm for collective decision-making. *Proceedings of the National Academy of Sciences*, 103(43):15906–15910, 2006.

[96] Stephen C Pratt, David JT Sumpter, Eamonn B Mallon, and Nigel R Franks. An agent-based model of collective nest choice by the ant Temnothorax albipennis. *Animal Behaviour*, 70(5):1023–1036, 2005.

[97] Matthew Rabin. The experimental study of social preferences. *Social Research: An International Quarterly*, 73(2):405–428, 2006.

[98] Werner Raub, Vincent Buskens, and Marcel ALM Van Assen. Micro-macro links and microfoundations in sociology. *The Journal of Mathematical Sociology*, 35(1-3):1–25, 2011.

[99] L. Rendell, R. Boyd, D. Cownden, M. Enquist, K. Eriksson, M. W. Feldman, L. Fogarty, S. Ghirlanda, T. Lillicrap, and K. N. Laland. Why Copy Others? Insights from the Social Learning Strategies Tournament. *Science*, 328(5975):208–213, 2010.

[100] Ariel Rubinstein. The electronic mail game: Strategic behavior under "almost common knowledge". *The American Economic Review*, pages 385–391, 1989.

[101] Andrey Rzhetsky, Jacob G Foster, Ian T Foster, and James A Evans. Choosing experiments to accelerate collective discovery. *Proceedings of the National Academy of Sciences*, 112(47):14569–14574, 2015.

[102] Matthew J. Salganik, Peter Sheridan Dodds, and Duncan J. Watts. Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, 311(5762):854–856, 2006.

[103] Thomas D. Seeley and Susannah C. Buhrman. Group decision making in swarms of honey bees. *Behavioral Ecology and Sociobiology*, 45(1):19–31, 1999.

[104] Ohad Shamir. Fundamental limits of online and distributed algorithms for statistical learning and estimation. In *Advances in Neural Information Processing Systems*, pages 163–171, 2014.

[105] Duncan Snidal. The game theory of international politics. *World Politics*, 38(1):25–57, 1985.

[106] Spyros Spyrou. Herding in financial markets: a review of the literature. *Review of Behavioural Finance*, 5(2):175–194, 2013.

[107] G. L. Stewart. A Meta-Analytic Review of Relationships Between Team Design Features and Team Performance. *Journal of Management*, 32(1):29–55, 2006.

[108] Peter Stone, Gal A Kaminka, Sarit Kraus, and Jeffrey S Rosenschein. Ad hoc autonomous agent teams: Collaboration without pre-coordination. In *Twenty-Fourth AAAI Conference on Artificial Intelligence*, 2010.

[109] Robert Sugden. Team reasoning and intentional cooperation for mutual benefit. *Journal of Social Ontology*, 1(1):143–166, 2015.

[110] James Surowiecki. *The Wisdom of Crowds*. Anchor, 2005.

[111] Richard S Sutton and Andrew G Barto. *Reinforcement Learning: An Introduction*. MIT Press, 1998.

[112] Georg Theiner, Colin Allen, and Robert L Goldstone. Recognizing group cognition. *Cognitive Systems Research*, 11(4):378–395, 2010.

[113] KA Thomas, P DeScioli, OS Haque, and S Pinker. The psychology of coordination and common knowledge. *Journal of Personality and Social Psychology*, 107(4):657–676, 2014.

[114] William R Thompson. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, pages 285–294, 1933.

[115] Michael Tomasello. *A Natural History of Human Thinking*. Harvard University Press, 2014.

[116] Michael Tomasello, Malinda Carpenter, Josep Call, Tanya Behne, and Henrike Moll. Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28(05):675–691, 2005.

[117] Wataru Toyokawa, Hye-rin Kim, and Tatsuya Kameda. Human collective intelligence under dual exploration-exploitation dilemmas. *PloS one*, 9(4):e95789, 2014.

[118] Arnout van de Rijt, Soong Moon Kang, Michael Restivo, and Akshay Patil. Field experiments of success-breeds-success dynamics. *Proceedings of the National Academy of Sciences*, 111(19):6934–6939, 2014.

[119] Sofía S Villar, Jack Bowden, and James Wason. Multi-armed bandit models for the optimal design of clinical trials: Benefits and challenges. *Statistical Science: A Review Journal of the Institute of Mathematical Statistics*, 30(2):199, 2015.

[120] Edward Vul, Noah Goodman, Thomas L Griffiths, and Joshua B Tenenbaum. One and done? Optimal decisions from very few samples. *Cognitive Science*, 38(4):599–637, 2014.

[121] Nir Vulkan. An economist's perspective on probability matching. *Journal of Economic Surveys*, 14(1):101–118, 2000.

[122] James P Walsh and Gerardo Rivera Ungson. Organizational memory. *Academy of Management Review*, 16(1):57–91, 1991.

[123] Ashley JW Ward, David JT Sumpter, Iain D. Couzin, Paul JB Hart, and Jens Krause. Quorum decision-making facilitates information transfer in fish shoals. *Proceedings of the National Academy of Sciences*, 105(19):6948–6953, 2008.

[124] Daniel M Wegner. Transactive memory: A contemporary analysis of the group mind. In *Theories of Group Behavior*, pages 185–208. Springer, 1987.

[125] Thomas N. Wisdom, Xianfeng Song, and Robert L. Goldstone. Social Learning Strategies in Networked Groups. *Cognitive Science*, 37(8):1383–1425, 2013.

[126] Anita Williams Woolley, Christopher F. Chabris, Alex Pentland, Nada Hashmi, and Thomas W. Malone. Evidence for a Collective Intelligence Factor in the Performance of Human Groups. *Science*, 330(6004):686–688, 2010.

[127] Wako Yoshida, Ray J Dolan, and Karl J Friston. Game theory of mind. *PLoS Computational Biology*, 4(12):e1000254–e1000254, 2008.

[128] Wu Youyou, David Stillwell, H Andrew Schwartz, and Michal Kosinski. Birds of a feather do flock together: Behavior-based personality-assessment method reveals personality similarity among couples and friends. *Psychological Science*, 28(3):276–284, 2017.