Bootstrapping an Imagined We for Cooperation

Ning Tang(ningtang@g.ucla.edu)
Stephanie Stacy(stephanieestacy@gmail.com)
MingLu Zhao(mz517@g.ucla.edu)
Gabriel Marquez(gabrielemarquez0@gmail.com)
Departments of Statistics, University of California, Los Angeles, USA
Tao Gao(tao.gao@stat.ucla.edu)
Departments of Statistics and Communication, University of California, Los Angeles, USA

Abstract

Remaining committed to a joint goal in the face of many enticing alternatives is challenging. Doing so while cooperating with others under uncertainty is even more so. Despite this, agents can successfully and robustly use bootstrapping to converge on a joint intention from randomness under the Imagined We framework. We demonstrate the power of this model in a real-time cooperative hunting task. Additionally, we run a suite of model experiments to answer some of the potential challenges to converging that this model could face under imperfect conditions. Specifically, we ask what happens when (1) there are increasingly many equivalent choices? (2) I only have an approximate model of you? and (3) my perception is noisy? We show through a set of model experiments that this framework is robust to all three of these manipulations.

Keywords: Theory of Mind; Bayesian inference; cooperation; shared agency

Introduction

How do you model an intention that lacks a mind? Or rather, one that exists — imperfectly — among multiple minds? Individual intentions and their definition have long occupied analytic philosophers, generating corresponding computational models of intentions that address how humans form intentions from beliefs and desires (Bratman, 1987). Despite the rich philosophical debate around their form, shared intentions have yet to receive that same rigorous treatment. Here, we provide such a formalized computational account of joint commitment. Buoyed by this philosophy, we believe our model ties together the evolutionary roots of collaboration with its modern empirical expressions. Unsurprisingly, creating such a model requires drawing on a number of different fields.

To understand the motivation behind this model, it is important to first discuss Gilbert's philosophy and her explicit definition of shared intentions. In her formulation, cooperating parties must create a joint commitment in order to share their intentions (Gilbert, 1999). They must intend to complete a task *as a body*. Such a commitment is not merely the sum of personal intentions to complete a task, but in a sense, a subordination of personal intentions to the shared one. This allows for partner regulation after shortcomings and requires consensus when commitments change. These consequences of Gilbert's formulation provide us with concrete, testable predictions that have appeared to varying degrees in empirical research.

In contrast to Gilbert, some philosophers have taken the stance that individual agency provides a sufficient framework for collaboration (Bratman, 2013). We do not suggest such a framework is incompatible with human cognition, but we feel Gilbert's joint commitments offered a more robust structure to implement.

In addition to deep theoretical support from philosophy, joint commitments have empirical credence from research in developmental psychology. Three phenomena in particular emerge in the behavior of young children (<5 years old) while engaged in joint commitments which match predictions from Gilbert's philosophy. First, when a partner breaks the commitment, children attempt to re-engage (Gräfenhain, Behne, Carpenter, & Tomasello, 2009). This indicates cooperation, even at an early age, involves reciprocal expectations of and obligations toward one's partner. Joint commitment can be thought of as an implicit social contract; thus, it is natural to have a mechanism for regulation, ensuring cooperation is robust.

Second, when children break the commitment, they acknowledge doing so (Gräfenhain, Behne, Carpenter, & Tomasello, 2009). While engaged in an activity with a partner, children were more likely to offer some conspicuous sign that they were leaving the activity if both partner and child had mutually agreed to engage in the activity in the first place. This distinguishes merely "doing the same thing" from truly sharing agency.

Third, children continue engaging in tasks until all partners are rewarded even when they have already received their share (Warneken, Chen, & Tomasello, 2006). To be clear, these actions went against the children's immediate personal utility, but when measured jointly, contributed to the utility of the group. Again, commitment to the shared intention predicts that individuals ought to display such behavior.

While this research in developmental psychology has contributed to the understanding and formalizing of infant minds, it has not yet led to computational models of a joint mind as we propose in this paper. To our knowledge, no research has proposed a formal computational model of shared intentions as we do here, though it is important to note that work has been done to model multi-agent collaboration, including collaborative norm building (Ho et al., 2016). We believe the lack of such a model may stem from the perceived illogical nature of separate individuals "sharing" a mind.

Nonetheless, we suggest that collaborators *imagine* such a joint mind — an analog of Gilbert's joint commitment — in order to engage in collaborative tasks. Research into social contracts supports the idea that cooperators view their collaboration from a "bird's eye perspective", where all individuals are reasoned about as a whole (Carpenter, Tomasello, & Striano, 2005). By modeling a controller with this perspective, what we refer to as the "Imagined We", we can offer a valid structure for shared agency in human collaboration that addresses the reality that human minds are private. In this paper, we formalize that model and present its performance on a collective hunting task.

Cooperative Hunting Task

To test our model, we adapt a previously developed non-cooperative hunting task for use in a cooperative environment (Gao, Newman, & Scholl, 2009). This task lies at the border between proposed evolutionary demands for cooperation and empirical studies of the same, exploring a current gap between the two. That is, modern empirical studies (such as the developmental psychology studies discussed earlier) cannot easily create the conditions which theorists propose led to early human collaboration. We believe that computational modeling allows for better exploration of those conditions and that a hunting environment mimics the broad strokes of early human collaboration.

We populate the environment (Fig. 1) with two hunters (also referred to as wolves) and at least two hunting targets (also referred to as sheep). Wolves aim to successfully catch the sheep while sheep aim to avoid the wolves. Agents in the environment can take one of nine actions at a given time-step {*move in any of the four cardinal directions or the four diagonal directions or stay in place*} in order to achieve their respective goals. The sheep move faster than the wolves, which requires wolves to collaborate by persistently chasing a single target. However, they have no predetermined target. Instead, they must come to a collective decision about which sheep to prioritize, which is accomplished using our model of shared agency.

In this task, wolves do not possess a mechanism for explicit communication. This is motivated by a prominent theory on the evolutionary origins of communication. It is believed that communication may only emerge in an environment where collaboration already exists (Tomasello, 2010).Otherwise, there is no adaptive advantage for developing a method of communication. Thus, any viable model of collaboration must first succeed in a scenario that lacks communication, and we have built our collective hunting task on that assumption.

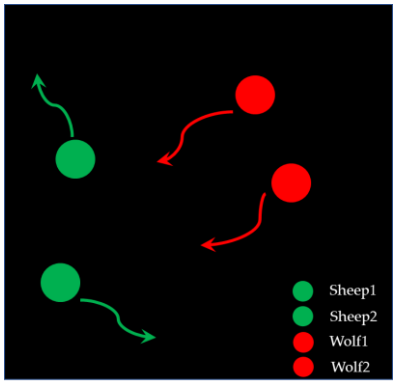

Figure 1: Cooperative Hunting Task.

Task performance is evaluated through achieved rewards. The wolves receive a joint reward (+1) upon the successful capture of either sheep. Each wolf also incurs a small negative reward (-0.01) at every time step to encourage faster chasing. Accumulated reward at the end of each trial is used as a dependent measure of the model's performance.

Though outside the aim of this paper, the psychophysics of perceiving *non*-cooperative chasing has been systematically studied in the field of perceived animacy (Gao, Newman, & Scholl, 2009; Gao, Scholl, & McCarthy, 2012). While we only report modeling results of this cooperative chasing task here, we are confident that this task can inspire future psychophysics work beyond this model.

Demos of our task and model can be found at: https://www.youtube.com/playlist?list=PL7v_qAmAikjzYia-dL0bPB3FCSerqTyC6

Bootstrapping Imagined We Framework

Here, we introduce a precise formulation of our model, including its computational foundation and its novel approach to shared agency. The primary question our model addresses is how shared agency can emerge and be maintained in human collaboration despite a changing environment and without explicit communication.

Our shared agency model builds on top of Theory of Mind (ToM) for individual agents. First, we explore previous work using ToM to model individual action planning and inference. Using that as a foundation, we introduce the Imagined We super agent in order to accommodate collaboration. This Imagined We is a reflection of joint commitment that allows the wolves to "espous[e] a goal as a body" (Gilbert, 2013). Of course, because this model lacks explicit communication, the wolves cannot "speak" to each other to create a single, unified version of this agent. Moreover, even agents that could communicate would find that signaling inaccuracies prevent them from creating that unified super agent instantly.

Instead, we propose a novel bootstrapping method wherein successive inference creates *distinct* super agents, unique to each individual agent, that converge over time to the same values. This method creates the Imagined We, which we will define more rigorously below.

Theory of Mind

The computational foundation of the model builds on ToM modeling work. ToM uses social reasoning to characterize the mind via a set of mental states — beliefs, desires, and intentions. These latent states define the ontology of mind. Beliefs are the informational states of the mind, desires are the motivational states of the mind, and intentions are the deliberative states of the mind (Bratman, 1987). For example,

someone walks by a $20 bill on the ground without picking it up. This can be explained in terms of your mental states: you didn't see it (beliefs), you didn't want it (desires), or you wanted it but were already committed to something else and didn't have the time to stop (intentions). This enumerative approach to mental states allows for computational accounts of action planning.

Action planning using ToM follows the "principle of rationality." Agents are assumed to plan actions that maximize their utility while minimizing their costs, all with respect to their underlying mental states. Agents select an action in a manner equivalent to sampling their available actions from a soft-max function typically used for approximately rational decision making, shown in Eq. 1. The parameter β controls how rational we believe the agent to be.

$$P(Action|Mind) \propto e^{\beta \mathrm{E}[U(action,mind)]} \quad (1)$$

Rationality provides a mechanism for action selection given the state of one's own mind, but it also provides observers a mechanism to reason about a mind given a set of actions. The reverse process of action selection, what is known as inverse planning, is an observer's Bayesian inference to figure out the most likely mind generating a set of observed actions in the environment (see Eq. 2).

$$P(Mind|Action, Environment) \propto P(Action|Mind, Environment)P(Mind|Environment) \quad (2)$$

This ToM inference framework has been successfully used to infer physical goals (Baker, Saxe, & Tenenbaum, 2009), social goals (Ullman et al., 2009), and joint beliefs and desires (Baker, Jara-Ettinger, Saxe, & Tenenbaum, 2017) from observed actions. Furthermore, inverse planning models have also been used to show how children make inferences about beliefs and desires to explain a variety of their behavior (Jara-Ettinger, Gweon, Schulz, & Tenenbaum, 2016; Jara-Ettinger, Floyd, Huey, Tenenbaum, & Schulz, 2019). In our model, we go beyond existing accounts by using ToM to model a joint mind.

Multi-Agent ToM: Bootstrapping Imagined We

Let's presume for a moment that a truly joint mind — rather than the Imagined We that we propose — governed shared agency. This "We" would be a super agent with its own mind containing beliefs, desires, and intentions. Using those mental states, it could rationally control the actions of agents, just as a person might rationally control their own hands. Its state and action space would simply be the joint state and action space produced by concatenating the individual agents' state and action spaces. And assuming the "We" agent governed shared agency, the contents of its mental states might be inferred from the actions of the jointly committed agents just as they could be inferred for a single agent using ToM. Now, understanding that no joint mind actually exists to control these agents, let's consider our proposal, the Imagined We.

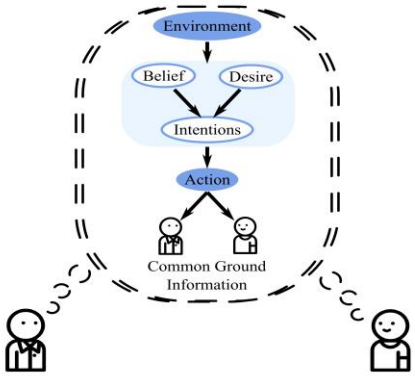

Figure 2. Imagined We Representation.

While similar in many ways to a real "We" agent, the Imagined We (Fig. 2) presents a unique distinction from standard ToM modeling. The Imagined We is, indeed, imagined. In reality, there is no shared mind to infer. Instead, each collaborating agent infers its *own version* of the Imagined We from its actions and its partner's actions in the shared environment. The Imagined We exists only as an inferred distribution of mental states that is unique to each collaborating agent.

Since there is no ground truth of "We" to infer, all agents can only reach agreement through bootstrapping (Fig. 3), agreement being achieved when the mental states of each agent's Imagined We align with the other agent's. Essentially, this is the process of determining what "We" want to do by looking at what "We" have done. We model the convergence of the Imagined We with three steps of computation. The Imagined We is designed to generically handle different types of uncertainty in latent mental states; however, in this collective hunting task, we only face uncertainty in joint intentions: which sheep is the joint goal.

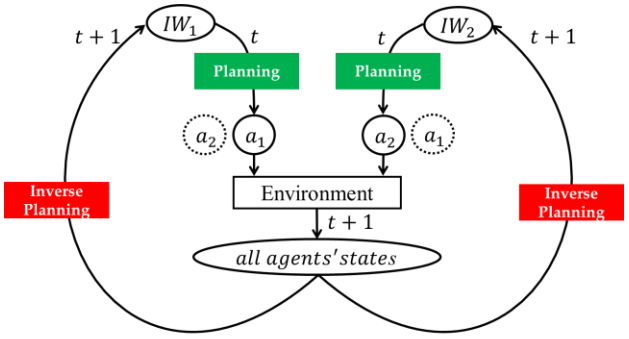

Figure 3. Bootstrapping Imagined We. The "IW" nodes represent the unique inferred distributions of mental states for each agent's Imagined We. The "a" nodes are the actions chosen given those inferred distributions, with the solid nodes being the actions each agent will actually take and the dashed nodes being the expected actions of each agent's partner. These actions are then observed by both agents and are used by each agent to update its Imagined We for the next time step.

(1) Goal Sampling: Each agent simply samples one sheep from its own goal distribution to pursue as the goal. This sheep becomes its target, and the agent proceeds by expecting the other agent will target the same sheep.

$$I_{xt} \sim P(IW_{xt}) \tag{3}$$

(2) Planning: Given a goal, each agent forms a plan of how all agents should pursue that goal rationally. The output of this planning process includes its own action to take, as well as an expectation of other agents' actions. This is essentially a centralized planning process.

We implemented this rational planning by combining online model-based simulation and off-line deep-reinforcement training, a framework inspired by Alpha-zero (Silver et al., 2018). Known as Monte Carlo Tree Search, the model-based simulation involves agents making predictions several time-steps in the future given their knowledge of the other agents' intentions. At every time step, the agent balances choosing actions it currently believes are the most rewarding and choosing actions that have gone unexplored. In the figure depicting this simulation (Fig. 4), Combining this simulation with the offline learning produces a policy as output, defining the probability of joint actions conditioning on the current state $\pi(S_{1t,2t,I_{xt}t})$.

Importantly, this rational planning phase does not imply human cognition necessarily uses the simulation and off-policy learning we utilize here. Rather, we assert that humans generally make ration plans, and in order to justify the rational inference of step 3, we ensure the agents act rationally with the planning engine described.

(3) Inference: After taking one's own action based on the policy determined in the planning phase, each agent observes the actions actually taken by other agents. This enables a

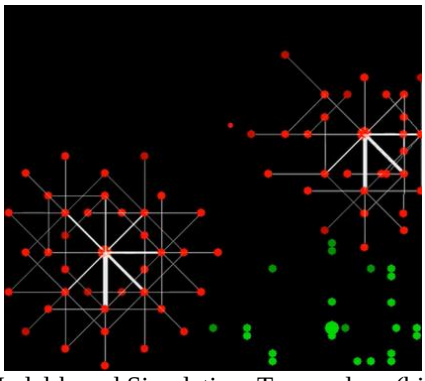

Figure 4: Model-based Simulation. Two wolves (big red circles) pursuing a single sheep (big green circles) at single time step. The model simulates multiple futures, going several steps into each. Smaller red and green circles indicate possible future locations for the agents. The line thickness and circle shade indicate how often a given action has been taken in simulation. The best action is the one taken most frequently.

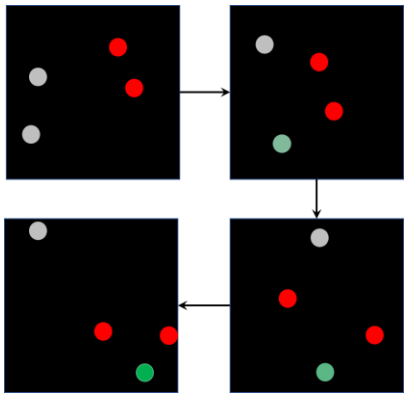

Figure 5: Convergence of Imagined We. The wolves (in red) initially infer that both sheep (in grey-green) are equally likely to be their joint goal. The exact shade of each sheep along the grey-green gradient represents how likely both wolves are to believe that a given sheep is that joint goal. In each successive time step, the wolves converge on the lower sheep, with this convergence visible in the shade change of that sheep as well as in the movements of the wolves

Bayesian ToM inference process: conditioning on the observed actions, each wolf computes the posterior probability of a given sheep being their joint goal.

$$P(IW_{x(t+1)}|IW_{xt}, A_{1t,2t}) \propto P(IW_{xt})P(A_{1t,2t}|IW_{xt}) \tag{4}$$

After updating the posterior of the Imagined We mind, each agent goes back to step (1), sampling a new goal and repeating the process. In Figure 5, we show the repeated implementation of these 3 computational steps as the Imagined We minds converge on one sheep as the chosen goal.

Modeling Experiments

Overview

Bootstrapping an Imagined We is potentially noisy and faces the challenge of convergence, particularly under imperfect conditions and high uncertainty. Here we report three modeling experiments, each challenging the robustness of the Imagined We model in a distinctive way that is inspired by human collaborative challenges. With these tests, we hope to demonstrate both the robustness of this model computationally and the validity of the Imagined We as a potential explanation of shared agency. Due to the stochastic nature of the simulations, we use accumulated reward as a dependent of performance.

Expt. 1: Multiple Alternative Targets

Human collaboration often involves a choice between pursuing multiple equivalent goals. Our first test introduces an increasing number of alternative targets for the wolves to assess the model's performance ability to handle this common real-

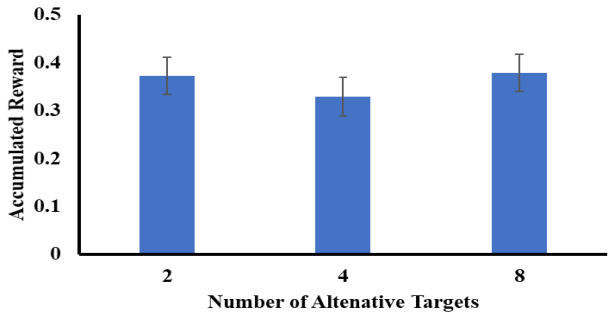

Figure 6: Results of Experiment 1.

world scenario. Presumably, with an increasing number of sheep, the two wolves might experience greater difficulty in choosing which sheep to pursue persistently.

We tested how well wolves were able to cooperate using joint commitment in the presence of 2, 4, and 8 equivalent goals, with 200 trials for each condition.

Results The performance for each set size condition is depicted in Figure 6. The number of alternative targets is not significant ($F(2, 597) = 0.466$, $p = 0.628$). These results reveal that, even in the presence of increasingly many equivalent options, the Imagined We Model achieves effective cooperative chasing. The agents converge, reaching a consensus through inference about We.

Expt. 2: Model Precision

Another common problem for collaboration stems from imprecise models of other agents. Collaborators do not always know the exact capabilities of their partners. Here we manipulate the precision of each agent's representation of the other agent's action space.

While each agent's own action is still selected from a set of 9, here they use simplified models of the other agent with a smaller action space. A nearest neighbor approach is adopted to map the real action to the action perceived by the other agent. The set size of the perceived action space is selected from 2, 3, 5, and 9 with 200 trials in each condition.

$$A'_{xt} = argmin \|A_{xt} - A\|^2, \quad A \in \mathcal{A}' \qquad (5)$$

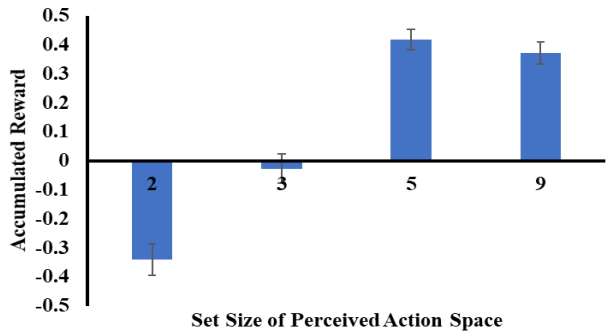

Figure 7: Results of Experiment 2.

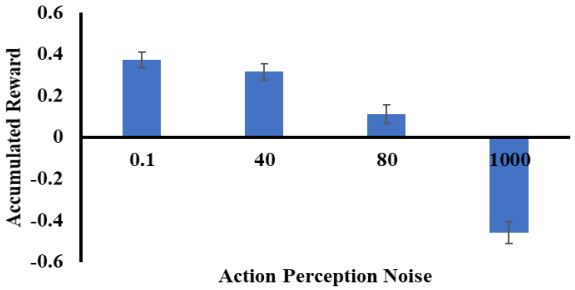

Figure 8: Results of Experiment 3.

Results Performance as a function of perceived action precision is shown in Figure 7. One-way ANOVA results reveal a significant main effect of action space size ($F(3, 796) = 62.91$, $p < .001$). Specifically, the accumulated rewards collected by the agents in the 5 and 9 action space conditions are significantly higher than in the 3 action space condition ($t(398) = 7.297$, $p < 0.001$; $t(398) = 6.320$, $p < 0.001$), which is significantly higher than in the 2 action space condition ($t(398) = 4.233$, $p < 0.001$).

These results revealed that the Imagined We does not require a perfectly precise model of other agents in order to converge on a shared target. Cutting the perceived action space nearly in half does not impact performance. However, more simplified action representations with fewer than 5 actions do significantly reduce the model's performance.

Expt. 3: Noisy Action Perception

Finally, we test the robustness of the Imagined We by introducing random noise in the agent's perception. This condition mimics human perceptual errors, which are another source of complications in collaborative action. A Gaussian noise is added to each agent's perception of the others' actions. Across trials, the variance of the Gaussian noise is selected from 0.1, 40, 80, and 1000, with 200 trials in each condition.

$$A''_{xt} \sim N\left(A_{xt}, \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}\right) \qquad (6)$$

Results Model performance as a function of perception noise is shown in Figure 8. One-way ANOVA results reveal a significant main effect ($F(3, 796) = 96.034$, $p < .001$). Specifically, the accumulated rewards collected by the agents in the 0.1, 40, 80 noise condition are significantly higher than than in the 1000 noise condition ($t(398) = 12.822$, $p < 0.001$; $t(398) = 11.895$, $p < 0.001$, $t(398) = 13.401$, $p < 0.001$). These results demonstrate that the Imagined We can tolerate a moderate amount of perceptual noise and only suffers only with a large amount of noise.

Conclusion

Inspired by philosophical and developmental studies of shared agency, we develop an Imagined We model and test it

in a multi-agent cooperative hunting task. The most important discovery is that it consistently converges under a variety of conditions, as the wolves iteratively come to an agreement on which sheep to jointly pursue. This model is relatively robust, performing well with a large number of potential targets, a reduced perceived action space, and a moderate amount of perceptual noise. Our study illustrates the rich potential of modeling human-like cooperative intelligence based on insights from developmental studies and analytic philosophy.

One additional finding concerns the lack of explicit communication in our model. In the designed task, agents had no method of communication other than the communication implicit within their movements. The model's success despite this fact emphasizes one of our underlying assumptions - that models of collaboration ought to be possible without communication. This lends further credence to the idea that collaboration predates communication from an evolutionary perspective by demonstrating that, at a minimum, collaboration in this multi-agent model can exist without communication.

Finally, though we chose to draw inspiration for the technical aspects of our model from the philosophy of Margaret Gilbert, there are other theories on shared agency. One prominent theory stems from the work of Bratman (2013), who uses an existing single agent framework to explain human collaboration at the small scale. Though the developmental psychology research we cited earlier supports human collaboration through the lens of a super agent (a la the theories of Gilbert), we believe future research should explore alternative computational accounts of theories on cooperation as other candidates to explain human cognition during collaboration.

Acknowledgments

This research was supported by DARPA PA 19-03-01 and ONR MURI project N00014-16-1-2007 to TG.

References

Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, *1*(4), 1-10.

Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition*, *113*(3), 329-349

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

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

Carpenter, M., Tomasello, M., & Striano, T. (2005). Role reversal imitation and language in typically developing infants and children with autism. *Infancy*, *8*(3), 253-278.

Gao, T., Newman, G. E., & Scholl, B. J. (2009). The psychophysics of chasing: A case study in the perception of animacy. *Cognitive psychology*, *59*(2), 154-179.

Gao, T., Scholl, B. J., & McCarthy, G. (2012). Dissociating the detection of intentionality from animacy in the right posterior superior temporal sulcus. *Journal of Neuroscience*, *32*(41), 14276-14280.

Ho, M. K., MacGlashan, J., Greenwald, A., Littman, M. L., Hilliard, E. M., Trimbach, C., Brawner, S., Tenenbaum, J. B., Kleiman-Weiner, M. & Austerweil, J. L. (2016). Feature-based joint planning and norm learning in collaborative games. In A. Papafragou, D. Grodner, D. Mirman, & J. C. Trueswell (Eds.), *Proceedings of CogSci 2016: Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 901-906). Austin, Texas: Cognitive Science Society.

Gilbert, M. (1999). Obligation and joint commitment. *Utilitas*, *11*(2), 143-163.

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

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

Jara-Ettinger, J., Floyd, S., Huey, H., Tenenbaum, J. B., & Schulz, L. E. (2019). Social pragmatics: Preschoolers rely on commonsense psychology to resolve referential underspecification. *Child dev.*

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

Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T. and Lillicrap, T. (2018). A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science*, *362*(6419), 1140-1144.

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

Ullman, T., Baker, C., Macindoe, O., Evans, O., Goodman, N., & Tenenbaum, J. B. (2009). Help or hinder: Bayesian models of social goal inference. In *Advances in neural information processing systems 22* (pp. 1874–1882).

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