

I'd do anything for a cookie (but I won't do that):* Children's understanding of the costs and rewards underlying rational action

Julian Jara-Ettinger (jjara@mit.edu), Hyowon Gweon (hyora@mit.edu), Joshua B. Tenenbaum (jbt@mit.edu), & Laura E. Schulz (lschulz@mit.edu)

Department of Brain and Cognitive Sciences
Massachusetts Institute of Technology, Cambridge, MA 02139 USA

Abstract

Humans explain and predict other agents' behavior using mental state concepts, such as beliefs and desires. Computational and developmental evidence suggest that such inferences are enabled by a principle of rational action: the expectation that agents act efficiently, within situational constraints, to achieve their goals. Here we propose that the expectation of rational action is instantiated by a *naïve utility calculus* sensitive to both agent-constant and agent-specific aspects of costs and rewards associated with actions. We show that children can infer unobservable aspects of costs (differences in agents' competence) from information about subjective differences in rewards (i.e., agents' preferences) and vice versa. Moreover, children can design informative interventions on both objects and agents to infer unobservable constraints on agents' actions.

Keywords: Naïve Utility Calculus; Social Cognition; Theory of Mind

Introduction

One of the assumptions underlying our ability to draw rich inferences from sparse data is that agents act rationally. In its simplest form, this amounts to the expectation that agents will take the shortest path to a goal subject to physical constraints imposed by the world (Gergely & Csibra, 2003). Even this simple formulation is inferentially powerful, supporting predictions about future events and inferences about unobserved aspects of events. For instance, if Sally hops over a wall to get a cookie, we assume that she would not hop, but walk straight to the cookie, if the wall weren't there. Studies suggest that even infants expect agents to act rationally. Infants can use information about an agent's goal and situational constraints (e.g., gaps, occluders, walls, etc.) to predict her actions (Gergely, Nádasdy, Csibra, & Bíró, 1995); an agent's actions and situational constraints to infer her goals (Csibra, Biro, Koos, & Gergely, 2003), and an agent's actions and goals to infer unobserved situational constraints (see Csibra et al., 2003 for review; see also Brandone & Wellman, 2009; Gergeley, Bekkering, & Kiraly, 2002; Phillips & Wellman, 2005; Schwier, Van Maanen, Carpenter, & Tomasello, 2006; Scott & Baillargeon, 2013).

Computationally, this approach to action understanding can be formalized as Bayesian inference over a model of rational action planning, such as a Markov Decision Process (MDP) (Baker, Saxe, & Tenenbaum, 2009, 2011; Ullman, Baker, Macindoe, Evans, Goodman, &

Tenenbaum, 2010; Jara-Ettinger, Baker, & Tenenbaum, 2012). MDPs are a framework widely used in artificial intelligence and other engineering fields for determining sequences of actions, or plans, an agent can take to achieve the highest-utility states in the most efficient manner, given a specification of the possible world states, the agent's possible actions and their likely outcomes, and the agent's utility function² (positively and negatively valued rewards) associated with different combinations of actions and world states. Bayesian inference over these probabilistic generative models can implement a form of rational *inverse planning*, working backwards from observations of an agent's actions to infer aspects of the agent's world model or utility function. Bayesian inverse planning accounts have been used to make fine-grained quantitative predictions of adults' judgments about an agent's desires, beliefs, and states of the world (Baker, et al., 2009; Baker, Saxe, & Tenenbaum 2011; Jara-Ettinger et al., 2012).

The details of this computational approach are not critical here, but it is helpful to consider the qualitative intuitions behind these models, as well as what they leave out, because they motivate our present work. Intuitively we can think of an agent's utility function as the difference between two terms: a (positive) reward term associated with goals to be achieved, measuring the value of a goal to the agent, and a (negative) cost term associated with actions that can be taken to achieve these goals, measuring the difficulty of an action. Formally, we can decompose the utility function (normally a joint function of the agent's state and actions) into a reward associated with each state, and a cost associated with each action:

$$U(a,s) = R(s) - C(a).$$

Note that observing an agent taking an action, a , to achieve state, s , implies only that the relative reward for s is significantly higher than the cost of a ; it does not determine either of these values in absolute terms: positing that the action has high cost but the goal state generates very high rewards, or that the action is relatively lower cost and the goal state is comparably lower in reward, maybe equally viable explanations of the same behavior. Psychologically however, high cost/high reward plans are very different from low cost/low reward ones. If Sally jumps over a wall to get a cookie is it because she likes the cookies so much

² In the artificial intelligence literature this is sometimes referred to as the *reward function*. However, since this function is derived from rewards minus costs, we refer to it as the *utility function* for clarity.

* Or That's the way the utility crumbles.

that the cost of climbing the wall is worth it, or because the obstacle is so trivial that it is worth surmounting even for a relatively mediocre cookie? Knowing the difference between these two scenarios is critical to understanding Sally's capabilities and motivations and predicting her future behavior.

Note also that both formal and informal accounts of rational action (Baker et al., 2009; Csibra et al., 2003) have assumed fixed, non-zero cost of actions determined by the structure of the environment (the distance to goals, the height of obstacles, etc.). Intuitively however, even in a constant environment, agents do not experience identical costs. There are both agent-independent (i.e., external, objectively observable) and agent-specific (i.e., internal and subjective) components to costs and rewards. Jumping over a high wall may always be more costly than jumping over a low one, however some people find jumping harder than others; thus the same wall exacts higher costs for some people than others. Similarly, getting two cookies may always be better than getting just one, but some people like cookies better than others; thus agents can obtain different rewards from the same number of cookies.

Such intuitions motivate an account of rational action that considers not just those aspects of the event that are constant across agents, but also those that vary between them: not just the height of the obstacle but Sally's *competence* to surmount it, and not just how many cookies Sally gets but how much Sally *values* them. We suggest that even very young children naturally understand agents' actions and goals in terms that go beyond a simple maximization of overall utility. Instead, children reason about the costs and rewards that form the utility function – an ability that we refer to as *naïve utility calculus* (See also Jara-Ettinger, Tenenbaum, & Schulz, 2013; Jara-Ettinger, Kim, Muentener, & Schulz, 2014).

Here we investigate three implications of a naïve utility calculus. First, children should understand that agents do not always pursue the states with the highest rewards because obtaining those states might also involve high costs; rational agents should maximize utilities rather than rewards. We test this understanding in Experiment 1 by looking at whether children can accurately infer an agent's subjective rewards (preferences) from the choices he makes by considering the relative costs of his choices. Second, children should understand that both preferences and competencies vary across agents, are not directly observable, and differ from situational constraints that uniformly affect all agents. In Experiment 2, we test this by introducing two agents who have different preferences but make identical choices; we look at whether children can use information about agents' preferences and choices to infer differences in their competence. Finally, children should be able to predict how changes in costs and rewards affect an agent's actions. Thus, they should be able to design interventions that render agents' choices informative with respect to their underlying competencies. We test this in Experiments 3 and 4.

Experiment 1

In Experiment 1 we look at whether children understand that an agent's choices depend on both the costs and the rewards associated with an action. In the test condition, children saw a puppet choose between two kinds of treats across two consecutive trials. In the first trial, both treats are equally costly to obtain and the puppet chose one of the two. In the second trial, the previously chosen treat was more costly to obtain, and the puppet chose the other treat (actual trial order counterbalanced). If children are insensitive to costs and assume the agent is acting only to maximize his rewards, they should conclude that the puppet likes both treats equally; he chooses each treat once. If, instead, children take costs into account and expect the puppet to maximize utilities, then children should infer that the puppet prefers the high-cost treat even though he only chose it on one of the trials.

Methods

Participants. 33 children (mean age: 5.85 years, range 5.0–6.9 years) were recruited at an urban children's museum³; children were assigned to a test condition or control condition. One participant was excluded from the test condition due to parental interference leaving $n = 16$ in each condition.

Stimuli. The stimuli consisted of a puppet (Ernie), a paper picture of a watermelon slice, a paper picture of a banana, and two cardboard boxes: a short box (30 cm high) and a tall box (62 cm high).

Procedure. Figure 1 shows the experimental setup. Participants were tested individually in a quiet room. The child and the experimenter sat on opposite sides of a small table where the tall and short cardboard boxes were placed. In the test condition, the experimenter introduced Ernie and then directed the child's attention to the two boxes. Participants were asked which box was the hardest for Ernie to climb. Children who chose the short box were corrected ($n = 5$). The experimenter then said, "It's easy for Ernie to climb the short box!" and had Ernie climb the short box swiftly and nod in agreement. Then the experimenter said, "It's hard for Ernie to climb the tall box. It makes him tired!" and had Ernie climb the tall box slowly, and running out of breath. Afterwards, the experimenter introduced the watermelon and the banana. The experimenter placed both treats on the short box. The experimenter had Ernie look at both treats and then choose the banana. The experimenter said, "When both treats are on the short box, Ernie always chooses the banana!" Next, the experimenter placed the watermelon on the short box and the banana on the tall box.

³ The choice of ages was motivated by pragmatic considerations of the experimental setup rather than developmental claims about the naïve utility calculus. Throughout we focus on five and six-year-olds because pilot data suggested that children of this age could handle the information-processing demands involved even in the hardest tasks (e.g., tracking different agents with different preferences or levels of competence, performing different actions in different contexts). We discuss this further in the General Discussion.

The experimenter had Ernie look at both treats and then choose the watermelon on the short box. The experimenter said, "When the watermelon is on the short box and the banana is all the way up on the tall box, Ernie always chooses the watermelon!" (Actual treat counterbalanced). The experimenter then placed both pictures on the table and asked, "Which treat does Ernie like the most?" Trial order and Ernie's preferred treat were counterbalanced throughout.

The control condition was designed to rule out the possibility that children might simply identify the preferred treat as the treat that moved locations between trials. The control condition was identical to the test condition except that on one trial both treats were placed on the table next to the short box and on the other trial one treat was placed on the table next to the short box and the other treat was placed next to the tall box. Because there was no difference in the costs associated with the two set-ups, we expected children to perform at chance in the control condition.

Results and Discussion

In the test condition, children were counted as succeeding on the task if they selected the treat that Ernie chose in the trial where both treats were equally costly to reach. Twelve of the sixteen children (75%) correctly selected Ernie's favorite treat ($p < 0.05$ by binomial test). See Figure 2. The results of the control condition suggest that these results were not due to children simply choosing the treat that moved locations. As expected, children in the control condition performed at chance (7 of 16 children (44%) chose the treat that Ernie chose when both treats were by the short box, $p = ns$ by binomial test).

Note that if the children expected Ernie to always pursue the treat with the highest reward, then their responses should have been equally split across the two treats in both conditions. However, even though Ernie chose both treats exactly once, children in the test condition successfully identified Ernie's preferred treat, suggesting they considered both his choices and the relative cost of those choices. These results suggest that children not only understand the external, objective costs of agent's actions (i.e., that a tall box is harder to climb than a shorter one) but can integrate this information with the agent's actions to infer unobservable mental states: the agent's subjective rewards, or preferences.

Experiment 2

In Experiment 1, only the external costs were manipulated. In Experiment 2 we look at whether children understand that the cost of an action can vary across agents and whether children can use information about agents' rewards to infer relative differences in agents' competence. In this task, children saw two puppets with different subjective rewards behaving identically. Based on this information, children were asked which of the two agents was more likely to have difficulty climbing. To succeed, children had to understand that both costs and rewards are agent-specific, and that agents act to maximize their utilities.

Methods

Participants. Thirty-six children (mean age: 5.8 years, range 5.0-6.9 years) were recruited from a children's museum and randomly assigned to either the Cookie-Cracker condition or the Clover-Daisy condition. Four children were excluded from analysis due to experimenter error ($N=2$), and parental interference ($N=2$), leaving a final sample of 16 children per condition.

Stimuli. A Cookie Monster puppet and a Grover puppet were used. A short cardboard box (20 cm high) and a tall cardboard box (51 cm high) were used for the puppets to climb. Paper cutouts of cookies and crackers or clover leaves and daisy flowers were used for the Cookie-Cracker and the Clover-Daisy conditions, respectively. We also used two additional pictures for the Clover-Daisy condition: one

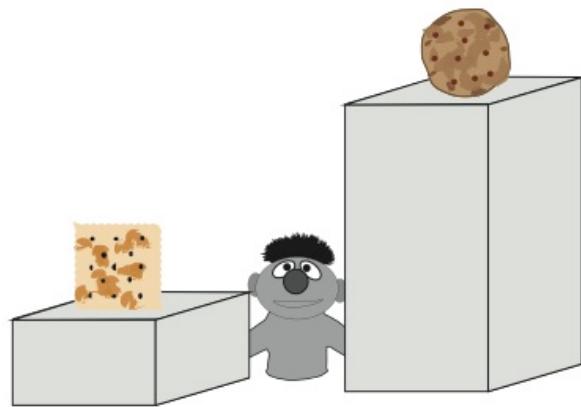


Figure 1. Example of experimental setup. All trials in all experiments consisted of a puppet choosing between two treats that could be placed either on the tall or the short box. We studied children's naïve utility calculus by varying the position of the objects, the puppet's choices, and the preference or competence information participants received.

of Grover surrounded by clovers and one of Cookie Monster surrounded by clovers and daisies.

Procedure. Participants were tested individually in a quiet room and sat across the table from the experimenter where the two boxes were set up. In the Cookie-Cracker condition, the experimenter showed the child paper cutouts of cookies and crackers and introduced the puppets. Children were told that Cookie Monster liked cookies better than crackers while Grover liked both treats equally (order counterbalanced). The preference information was repeated twice and children were prompted to ensure they remembered the information (e.g., "Remind me, does Cookie Monster like cookies? Yes, he loves cookies. And does he like crackers? Not so much."). Children who gave wrong answers were corrected. Next, children were told that both puppets could climb the short box, but the big box was so tall and hard to climb that only one of the puppets could climb up to the top. Children were told that in order to find out which puppet was the better climber we would place treats on the boxes and let the puppets choose a treat. In the first trial, a cracker and a cookie were placed on the short box. Each puppet approached the short box individually

(while the other puppet was absent), looked at both treats, and picked the cookie (order counterbalanced). In the second trial, the cracker was once again placed on the short box, but the cookie was now placed on the tall box. Once again, each puppet approached the boxes individually and looked at both treats, but this time both puppets picked the cracker. Children were then asked, "Which puppet do you think is the one who *cannot* climb?"

Because children might think that Cookie Monster could not climb for reasons irrelevant to the experiment (e.g., because cookie eaters are unhealthy), the Clover-Daisy condition was set up such that Grover was the puppet who couldn't climb. In this condition, Grover liked clovers better than daisies but Cookie Monster liked both equally. Although we chose clovers as the preferred stimuli for Grover hoping that children would easily associate the two (i.e., because Grover rhymes with clover), pilot data showed that children had a hard time remembering the puppets' preferences. Thus we added a picture of Grover with clovers and Cookie Monster with both clovers and daisies to help children remember the puppets' preferences. All other aspects of the two conditions were identical.

Results and Discussion

In both conditions, children successfully used the preference information to make competence judgments. In the Cookie-Cracker condition, 12 of the 16 children correctly identified Cookie Monster as the incompetent puppet ($p < 0.05$ by binomial test). In the Clover-Daisy condition, 13 out of the 16 children correctly identified Grover as the incompetent puppet ($p < 0.01$ by binomial test). See Figure 2.

Children's ability to distinguish agents' competences here is especially striking because both puppets behaved identically: each puppet chose each treat once, and neither climbed the tall box. In fact, neither puppet even attempted to climb the tall box. Instead they always *chose* to climb the small box, and always succeeded in their actions. In order for children to draw different conclusions about the competence of the two agents, children had to infer that the costs of climbing the tall box influenced the agents' choices. These results are consistent with our hypothesis that children evaluate agents through a naïve utility calculus that includes a principle of rational expectation.

Experiment 3

Experiments 1 and 2 suggest that children are able to represent and infer agent-specific competencies and preferences. In Experiment 3, we took a step further to investigate children's understanding of agent-independent (external) and agent-dependent (subjective) costs by asking whether children could manipulate the objective costs associated with different rewards so that an agent's actions would reveal his underlying competence.

Methods

Participants. Seventeen children (mean age: 6.0 years, range 5.1-6.8 years) were recruited at an urban children's museum and randomly assigned to either the Cookie-

Cracker stimuli (N=8) or the Clover-Daisy stimuli (N=8)⁴. One child failed to design an intervention and was therefore excluded from analysis.

Stimuli. The same stimuli used in Experiment 2 were used in Experiment 3.

Procedure. The experimenter first introduced the puppet to the child. Children given the Cookie-Cracker stimuli were told that Cookie Monster liked cookies better than crackers; children given the Clover-Daisy stimuli were told that Grover liked clovers better than daisies. The experimenter then said, "Here's a tall box, and here's a short box. It's very hard to climb the tall box, and we don't know if Cookie Monster (or Grover) can do it." She then gave the child two objects (a cookie and a cracker, or a clover and a daisy) and said, "We are going to put one of them on top of the tall box and the other on top of the little box. After that we are going to see what Cookie Monster does and see if he can climb. Where do you want to put them?"

Results and Discussion

As predicted, 14 of the 16 children made the informative intervention, putting the object with higher subjective reward in the more costly position ($p < 0.01$ by binomial test). See Figure 2. This suggests that children can predict how agents might act in the world as a function of the costs and rewards. They can then use this information to design interventions that are informative about agents' competence.

Although the task is very simple, it illustrates how combinations of costs and rewards could be (or fail to be) informative about unobservable properties of agents. In this task, children were asked to combine a high-reward (HR) and a low-reward (LR) object with a high-cost (HC) and a low-cost (LC) action to generate a utility function. Agent-independent knowledge of the costs tells us that climbing the tall box is always more costly than climbing the short box ($HC > LC$). However, the exact difference between these costs is unobservable and specific to each agent. The higher the agent's competence, the smaller the cost difference ($HC - LC$) is likely to be, but children do not know the absolute value of this quantity to begin with, just as they do not know the absolute difference $HR - LR$, only that $HR > LR$. If in the experiment we place the high reward object on the low-cost location, the agent can choose between a low-cost-high-reward plan ($HR - LC$), and a high-cost-low-reward plan ($LR - HC$). Here the agent's competence plays no role; it is always better to pick the high-reward object (because $HR - LC > LR - HC$ for any values of these quantities as long as $HR > LR$, $HC > LC$). Thus the choice between these two plans reveals nothing about the agent's competence.

If, instead, the high-reward object is placed in the high-cost location, then the agent's rational action choice becomes dependent on his competence. If the agent is very competent, then the difference between the high-cost plan and low-cost plan ($HC - LC$) is relatively small compared to

⁴ Children were arbitrarily assigned to one of the two sets of stimuli since the results of Experiment 2 suggested that there was no effect of stimulus set.

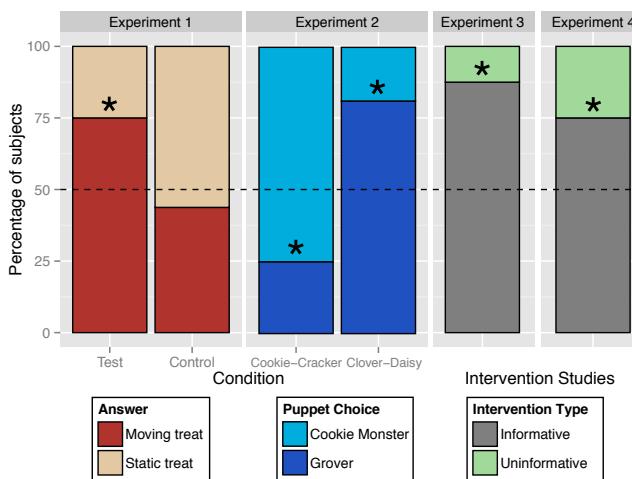


Figure 2. Results from all Experiments.

the difference in expected rewards ($HR - LR$); thus the high-reward-high-cost plan is likely to have a higher utility than the low-reward-low-cost plan ($HR - HC > LR - LC$). However, if the agent is less competent, then the difference between the high-cost plan and low-cost plan is relatively large ($HC - LC$) and the low-reward-low-cost plan becomes more likely to be the highest utility choice ($HR - HC < LR - LC$). Determining the informative intervention requires generating appropriate utility functions that depend on these agent-specific attributes.

Experiment 4

Experiment 3 suggests that children can selectively intervene on a desired object to infer an agent's competence. In Experiment 4, we look at whether children can selectively intervene on agents with different preferences to infer their competence (e.g., by picking the agent whose utility functions, given particular rewards and external constraints, will be informative about his subjective costs). Additionally, because children in Experiment 3 may have simply believed that more desirable objects should be placed in higher places (i.e., because parents often put treats out of children's reach), in Experiment 4 we had each treat be the favorite of one of the puppets.

Methods

Participants. Sixteen children (mean age: 6.0 years, range 5.0-6.9 years) were recruited at an urban children's museum.

Stimuli. The same stimuli used in Experiment 3 were used in Experiment 4.

Procedure. Experiment 4 began identically to the Cookie-Cracker condition in Experiment 2. The experimenter introduced Cookie Monster and Grover, the paper cookies and crackers, and the boxes. This time, Cookie Monster preferred cookies to crackers and Grover preferred crackers to cookies. As in Experiment 2, the experimenter told the child, "Both of our friends can climb up the small box. The big box is really hard to climb. One of our friends can climb it and one of our friends cannot. But we don't know which one can climb and which one cannot." The experimenter

then placed a cookie on the tall box and a cracker on the short box (object on tall box was counterbalanced). Children were asked, "If we want to figure out which of our friends can climb, which friend should we send in?"

Results and Discussion

We were interested in which puppet children chose to test. The intervention was considered informative if the child chose the puppet that preferred the treat on the tall box (i.e., cookies for Cookie Monster, crackers for Grover). Twelve of the 16 children made the informative intervention ($p < 0.05$ by binomial test). See Figure 2.

To succeed in this task, children had to predict how different agents would act as a function of their utilities, given common situational constraints. The agent whose preferred treat was on the short box had an uninformative utility function: he should always climb the short box no matter his competence (because $HR - LC > LR - HC$, using the notation of Experiment 3). By contrast, the agent whose preferred treat was on the tall box had an ambiguous utility function that could be resolved by his choice. If he were competent enough to climb the tall box easily (so that $HC - LC$ is relatively small, and thus $HR - HC > LR - LC$), he would be expected to climb to get his preferred treat. If he were not so competent (so that $HC - LC$ is large, and thus $LR - LC > HR - HC$), he would be expected to choose the less preferred treat on the short box. These results suggest that children can assign different sets of costs and rewards to agents under the same situational constraints and predict how the agents would act upon the resulting utilities.

General Discussion

The results of these studies suggest that young children understand how agents act in the world as a function of costs and rewards; we refer to the ability to engage in this kind of reasoning as a naïve utility calculus. Our findings suggest that children understand that there are unobservable, agent-specific aspects of costs and rewards, can make predictions about these unobservable variables, and can design informative interventions to infer them. Experiment 1 showed that children understand that agents act not to maximize rewards, but to maximize overall utility, such that agents will sometimes forego a high reward option because the costs of obtaining it are too high. Experiment 2 showed that children understand that competence constraints, unlike situational constraints, are agent-specific and cannot be directly observed; children were able to infer differences in agents' competence using information about their preferences, even given a constant environment in which agents engaged in identical actions. Experiments 3 and 4 showed that, in addition to being able to infer the components of utility functions, children can predict the behavior of agents with different costs and rewards, and thus can design interventions that are informative about agents' competence. Collectively, these studies suggest that children reason about agents' actions and goals in terms of utility functions, consistent with the idea that a naïve utility calculus underlies our social judgments even in early childhood.

These studies also raise several questions for further research. In each of these tasks, children simply had to distinguish two distinct preferences and two levels of competence (indeed, simply whether an agent was competent or not). We do not know to what extent children can use preference information to infer agent competence in more complex scenarios, involving graded preferences and graded levels of competence. Similarly, we do not know to what extent children can use competence information to infer graded levels of preferences. Finally, although our experiments suggest that children can infer rewards (Experiment 1) and costs (Experiment 2) when the other factor is fixed, we do not know whether children can jointly infer costs and rewards from situational constraints and observable actions.

As noted, the choice to test five and six-year-old children was a pragmatic one given the information-processing demands of the experimental designs. However, there is mounting evidence that humans engage in relatively rich psychological reasoning even as infants (e.g., Gergely & Csibra, 2003; Hamlin, Wynn, & Bloom, 2007; Onishi & Baillargeon, 2005). This suggests that a naïve utility calculus might play a role in children's inferences much earlier in development. Although we chose this age range for our initial investigation, further research might investigate the origin and developmental trajectory of these abilities.

Collectively, these studies test some of the fundamental assumptions of a naïve utility calculus, and look at whether children are sensitive to these principles even in early childhood. Children are not only sensitive to information about the costs and rewards of actions, but can also act on the world to learn about subjective components of these variables. This information supports rational inferences about agents' competencies even early in development, suggesting that a naïve utility calculus may lie at the heart of children's precocious social reasoning.

Acknowledgments

We thank the Boston Children's Museum and the families who volunteered to participate. We thank Eric Garr, Susie Sol Je Lee, Mika Maeda, Kristina Presing, Jessica Wass, and Jenny Yang for help with recruitment, piloting, and data collection. Special thanks to Nancy Kanwisher, Kristina Presing, Rebecca Saxe, Kim Scott, and Tomer Ullman for useful comments and discussions. This material is based upon work supported by the Center for Brains, Minds, and Machines (CBMM), funded by NSF STC awards CCF-1231216, and by the Simons Center for the Social Brain.

References

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

Baker, C. L., Saxe, R. R., & Tenenbaum, J. B. (2011). Bayesian theory of mind: Modeling joint belief-desire attribution. In *Proceedings of the thirty-second annual conference of the cognitive science society* (pp. 2469-2474).

Brandone, A. C., & Wellman, H. M. (2009). You Can't Always Get What You Want Infants Understand Failed Goal-Directed Actions. *Psychological Science*, 20(1), 85-91.

Csibra, G., Bíró, S., Koós, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science*, 27(1), 111-133.

Gergely, G., Bekkering, H., & Király, I. (2002). Developmental psychology: rational imitation in preverbal infants. *Nature*, 415(6873), 755-755.

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naïve theory of rational action. *Trends in cognitive sciences*, 7(7), 287-292.

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

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, 450(7169), 557-559.

Jara-Ettinger, J., Baker, C. L., & Tenenbaum, J. B. (2012). Learning what is where from social observations. In *Proceedings of the Thirty-Fourth Annual Conference of the Cognitive Science Society* (pp. 515-520).

Jara-Ettinger, J., Tenenbaum, J. B., & Schulz, L. E. (2013). Not so innocent: Reasoning about costs, competence, and culpability in very early childhood. In *Proceedings of the Thirty-Fifth Annual Conference of the Cognitive Science Society* (pp. 663-668).

Jara-Ettinger, J., Kim, N., Muentener, P., & Schulz, L. E., (2014). Running to do evil – How costs incurred by a perpetrator affect moral judgment. In *Proceedings of the Thirty-Sixth Annual Conference of the Cognitive Science Society*.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs?. *Science*, 208(5719), 255-258.

Phillips, A. T., & Wellman, H. M. (2005). Infants' understanding of object-directed action. *Cognition*, 98(2), 137-155.

Schwier, C., Van Maanen, C., Carpenter, M., & Tomasello, M. (2006). Rational imitation in 12-month-old infants. *Infancy*, 10(3), 303-311.

Scott, R. M., & Baillargeon, R. (2013). Do Infants Really Expect Agents to Act Efficiently? A Critical Test of the Rationality Principle. *Psychological science*, 24(4), 466-474.

Ullman, T. D., Baker, C. L., Macindoe, O., Evans, O. R., Goodman, N. D., Tenenbaum, J. B. (2010). Help or hinder: Bayesian models of social goal inference. *Neural Information Processing Systems* (Vol. 22, pp. 1874-1882).

Wellman, H. M., & Woolley, J. D. (1990). From simple desires to ordinary beliefs: The early development of everyday psychology. *Cognition*, 35(3), 245-275.

Yuill, N. (1984). Young children's coordination of motive and outcome in judgements of satisfaction and morality. *British Journal of Developmental Psychology*, 2(1), 73-81.