Children's understanding of the costs and rewards underlying rational action

Julian Jara-Ettinger [a], Hyowon Gweon [b], Joshua B. Tenenbaum [a], Laura E. Schulz [a,*]

[a] Department of Brain and Cognitive Sciences, MIT, Cambridge, MA 02139, United States
[b] Department of Psychology, Stanford University, Stanford, CA 94305, United States

A B S T R A C T

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. In four experiments, we show that, given an agent's choices, children (range: 5–6 year olds; $N = 96$) can infer unobservable aspects of costs (differences in agents' competence) from information about subjective differences in rewards (differences in agents' preferences) and vice versa. Moreover, children can design informative experiments on both objects and agents to infer unobservable constraints on agents' actions.

© 2015 Elsevier B.V. All rights reserved.

1. 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 were not 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 develop expectations about her actions (Gergely, Nádasdy, Csibra, & Bíró, 1995); an agent's actions and situational constraints to

form expectations about her goals (Csibra, Bíró, Koós, & Gergely, 2003), and an agent's actions and goals to reason about unobserved situational constraints (see Csibra et al., 2003 for review; see also Brandone & Wellman, 2009; Gergely, Bekkering, & Király, 2002; Phillips & Wellman, 2005; Schwier, Van Maanen, Carpenter, & Tomasello, 2006; Scott & Baillargeon, 2013).

Computationally, goal inference using the principle of rational action can be formalized as finding the agent's unobservable utility function (a function that describes the utility the agent receives in each possible state of the world).[1] Borrowing from frameworks widely used in artificial intelligence and other engineering fields (e.g., Markov Decision Processes), it is possible to determine the actions a rational (or efficient) agent should take to maximize a utility function, given the environmental constraints. Bayesian

---

[1] In the artificial intelligence literature this is sometimes referred to as the *reward function* (e.g., Bellman, 1957). However, since this function is derived from rewards minus costs, we refer to it as the *utility* function for clarity.

* Corresponding author.

inference over these probabilistic generative models forms a rational *inverse planning* mechanism, working backwards from the observed actions to infer the unobservable utility function the agent is maximizing. Bayesian inverse planning accounts have been used to make fine-grained quantitative predictions of adults' judgments about an agent's goals, beliefs, desires, and states of the world (Baker, Saxe, & Tenenbaum, 2009, 2011; Jara-Ettinger, Baker, & Tenenbaum, 2012).

Although the details of this computational approach are not critical here, 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 these actions. Formally, we can decompose the utility function into a reward associated with the goal state, and a cost associated with the necessary actions:

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

If we see an agent take actions a to reach state s, we can conclude that the reward for s is significantly higher than the cost of a; but we cannot determine their exact values. Consider Sally, who climbs over a wall to get a cookie. Sally's goal is clearly to get the cookie, and the reward for getting the cookie must be higher than the cost for climbing over the wall. However, we cannot tell if she chose to jump because she likes the cookies so much that the cost of climbing was well worth it (a high cost/high reward plan), or because the obstacle is so trivial that it is worth surmounting even for a relatively mediocre cookie (a low cost/low reward plan). Although the underlying costs and rewards are irrelevant when interpreting Sally's goal (getting a cookie), high cost/high reward plans are psychologically very different from low cost/low reward ones. If Sally incurred a high cost for the cookie, she will likely also try to get a cookie in other situations where the costs are lower. However, if Sally incurred a low cost for the cookie, she may forego the cookie when the cost increases. Intuitively, inferring Sally's utility tells us her goal, but understanding the underlying costs and rewards allow us to understand Sally's capabilities and motivations and to predict her future behavior.

Importantly, costs and rewards have both external and internal components. Some aspects of costs and rewards apply across agents: climbing over a high wall is always more costly than climbing over a low one, and two cookies are typically more rewarding than one. However, other aspects of costs and rewards differ across agents: Some people find climbing harder than others and some people like cookies better than others. 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 we naturally understand agents' actions and goals in terms that go beyond a simple maximization of the overall utilities. Instead, we 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, Kim, Muentener, & Schulz, 2014; Jara-Ettinger, Tenenbaum, & Schulz, 2013).

Note that this understanding of rational action requires more sophisticated reasoning than merely an understanding of goal-directed actions. As researchers have noted, reasoning about the goals of an agent's actions does not even necessarily require theory of mind. A learner could infer that actions are (or are not) efficient with respect to a goal and environmental constraint without imputing any mental states to agents (see Csibra & Gergely, 1998; Gergely & Csibra, 1997, 2003 on the "teleological stance" in understanding rational action). Such non-mentalist inferences may indeed underlie some successes at social cognition very early in infancy (e.g., Skerry, Carey, & Spelke, 2013; Sommerville, Woodward, & Needham, 2005). Researchers have similarly proposed that early studies of theory of mind (e.g., Onishi & Baillargeon, 2005; Southgate, Chevallier, & Csibra, 2010; Southgate, Senju, & Csibra, 2007) rely on implicit knowledge distinct from the explicit representations that emerge later in development (e.g., Perner & Roessler, 2012).

However, some aspects of the naïve utility calculus might require representations more sophisticated even than many tasks that clearly do require explicit theory of mind. We might predict, for instance, that children should come to understand not only that one agent likes something and another agent does not (see e.g., Repacholi & Gopnik, 1997) but that two agents can like the same thing to different degrees. Similarly, children should come to understand not only that one agent can perform an action and another cannot (Tomasello et al., 2005) but that two agents might perform (or fail to perform) the same action and yet incur, or expect, different costs (i.e., because one agent is more competent than another). Moreover, children should be able to infer differences in agents' subjective rewards from information about differences in their subjective costs (and vice versa), even in the absence of any explicit behavioral cues indicating that one agent is more motivated, or more competent than another, and in contexts where agents have identical epistemic access and face identical situational constraints.

In short, our study looks not at children's understanding of rational action or theory of mind in general, but at children's ability to infer unobservable individual differences among agents. To date, most research on children's reasoning about individual differences has occurred in the context of children's understanding of personality traits (e.g., that someone is "lazy" or "shy"). Research suggests that such trait understanding does not develop until seven or eight (Berndt & Heller, 1986; Kalish, 2002; Rholes & Ruble, 1984; Rotenberg, 1980, 1982; Ruble & Dweck, 1995) although more recent work suggests that it may emerge by age five (Liu, Gelman, & Wellman, 2007; Seiver, Gopnik, & Goodman, 2013). While our study does not require children to treat individual differences as enduring, stable traits, we do require children to reason about the unobservable, internal structure of goal-directed actions as a consequence of individual differences between

agents. Additionally, our study requires children to impute different mental states to agents given identical evidence about their behavior. The ability to understand that agents who perceive ambiguous evidence might interpret it differently is also a relatively late development (Carpendale & Chandler, 1996; Chandler & Helm, 1984). Thus here we focus our investigation on five and six-year-old children.

We investigate three implications of a naïve utility calculus. First, children should understand that costs influence an agent's choices. That is, 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 she makes by considering the relative costs of her choices. Second, children should understand that both rewards and costs 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. We ask whether children can manipulate the position of objects, or the role of agents, in order to gain information about agents' competencies. We test this in Experiments 3 and 4.

2. Experiment 1

In Experiment 1 we look at whether children understand that an agent's choices depend on the expected costs and rewards associated with an action. If children understand that agents act efficiently toward their goals but do not consider the associated costs and rewards, then they may fail to distinguish actions that maximize the rewards from actions that maximize the utility (the difference between costs and rewards). Suppose, for instance, that an agent reaches for a banana on one trial (when the banana is closer than a watermelon) and a watermelon on the next trial (when a watermelon and a banana are equidistant from the agent). If children simply believe that agents act to maximize rewards, then children might infer that the agent liked the watermelon and bananas equally because the agent engaged in efficient goal-directed actions on both trials. If, however, children consider both the costs and rewards of actions, they should instead infer that the agent prefers the treat chosen when the costs were equivalent: the watermelon.

We test exactly this in Experiment 1. Children see a puppet choose between two kinds of treats on two consecutive trials. In one trial (different cost trial), the cost of getting each treat is different, and the puppet chooses the low-cost treat. In another trial (same cost trial), the cost of obtaining each treat is matched and the puppet chooses the treat it had previously foregone (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 chooses it on only one of the trials.

2.1. Methods

2.1.1. Participants

32 children (mean age: 5.85 years, range 5.0–6.9 years) were recruited at an urban children's museum; one additional participant was tested but excluded from analysis and replaced due to interference from a sibling (see Section 2.2). Children were assigned to a test condition or control condition ($n = 16$ per condition).

2.1.2. 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).

2.1.3. Procedure

Fig. 1 shows the experimental setup. Participants were tested in a quiet room at the museum in the presence of their caregiver. 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. Afterward, 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. 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!" 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.

In our design, one treat was always placed on the low box, while the other moved from one box to the other. The control condition was designed to rule out the possibility that children might simply select the moving treat over the static treat one because the moving treat was more salient. The control condition followed the same logic as the test condition except that the treats were
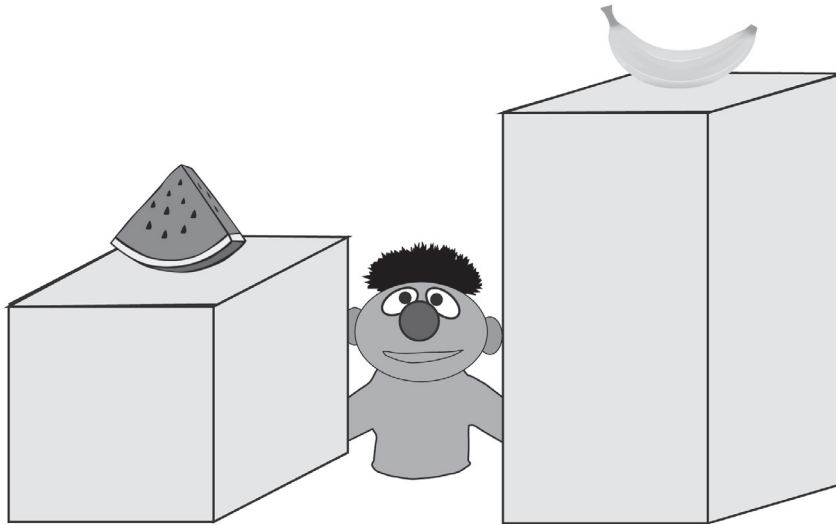
Fig. 1. Example of experimental setup. All trials in all experiments consisted of a puppet choosing between two treats that could be placed either on the tall or the short box. We studied children's naïve utility calculus by varying the position of the objects, the puppet's choices, and the preference or competence information participants received.

placed next to the boxes rather than on top and the experimenter substituted "next to" for "on" (e.g., "When both treats are next to the short box, Ernie always chooses the banana!") In this condition, both treats had equally low costs and the puppet chose each treat once; thus we expected children to perform at chance.

2.2. Results and discussion

All videotapes were coded by the first author for inclusion and children's responses to the test question; 100% of the videotapes were recoded on both measures by a second coder blind to hypotheses and conditions. The parents of two children did not consent to videotape and their responses were judged online. Children were excluded from analyses and replaced if the second coder judged A. that the items were not placed equidistant from the child or that the experimenter had otherwise cued the child's response (no children were excluded on these grounds) or B. if a parent or sibling interfered with the task (*n* = 1). 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. Intercoder agreement on all measures was 100%. Consistent with recent concerns about null hypothesis testing (e.g., Cohen, 1994; Cumming, 2013) we report confidence intervals throughout and report exact *p*-values as a secondary measure (reporting one-tailed tests when directional predictions warrant).

Twelve of the sixteen children correctly selected Ernie's favorite treat (75%; 95% CI: 56.25–100%); the remaining four children incorrectly selected the other treat. See Fig. 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

chose the treat that Ernie chose when both treats were by the short box (44%; 95% CI: 18.75–68.75%).

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 the relative cost of his choices. These results suggest that children not only understand the agent-invariant cost of 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.

3. Experiment 2

Experiment 1 suggests that children understand that agents maximize utilities and not rewards and thus costs influence agents' choices. However, this task only required children to understand the agent-invariant aspect of costs (i.e., that taller boxes are more costly to climb than shorter boxes). In Experiment 2 we look at whether children can use differences in agents' expected subjective rewards (their preferences) to infer agents' expected costs (their competencies).

In this task, children are introduced to two puppets. Each puppet can choose between a treat that is relatively costly to obtain and a treat that is relatively easy to obtain. One of the puppets likes both treats equally; the other puppet prefers the more costly treat. Children then see both puppets choose the less costly treat. Children are asked which puppet is unable to perform the high cost action. Critically, neither puppet ever attempts the costly action. Thus, in contrast to previous work (e.g., Tomasello et al., 2005) there are no behavioral cues to indicate whether
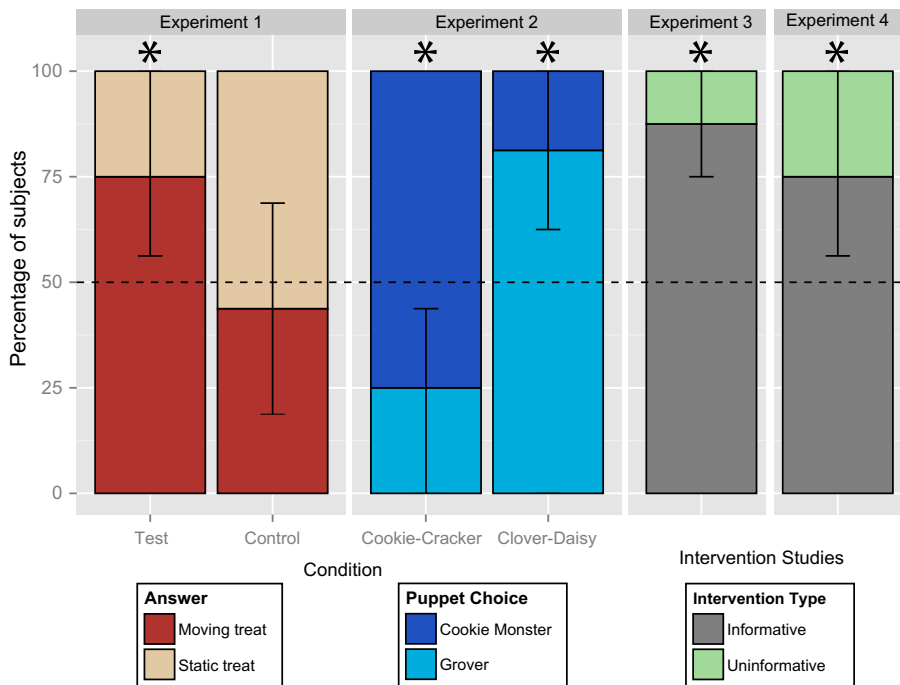
Fig. 2. Results from all experiments. Each bar shows the distribution of participant's responses across the four experiments. Vertical black lines show 95% confidence intervals bootstrapped from the data.

agents are unwilling or unable to perform the action. However, if an agent assigns identical reward to both treats, he should never attempt the more costly action insofar as he is maximizing his utilities. By contrast, if an agent assigns a high reward to the outcome associated with the high cost action, he should attempt the action unless the costs exceed the reward. Thus the second agent's actions are more likely to be informative about the agent's subjective costs than the actions of the first agent. If children are able to infer expected costs based on expected rewards, they should infer that the puppet with the preference had difficulty performing the action.

3.1. Methods

3.1.1. Participants

Thirty-two 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 ($N = 16$ in each condition). Four additional children were tested but excluded from analysis and replaced due to experimenter error ($N = 2$) and interference from siblings ($N = 2$).

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

3.1.3. Procedure

Participants were tested in a quiet room at the museum in the presence of their caregiver. Children sat across the table from the experimenter. The two boxes were on the table. 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 ($N = 3$). Next, children were told that both puppets could climb the short box, but the tall box was so 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 could not 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.

3.2. Results and discussion

All videotapes were coded blind to condition by the first author for inclusion and children's responses to the test question; 62% of the videotapes were recoded on both measures by a second coder blind to hypotheses and conditions. The parents of two children did not consent to videotape and their responses were judged online. Intercoder agreement was 100%. 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 (75%; 95% CI: 56.25–100%). The remaining four children identified Grover as the incompetent puppet. In the Clover–Daisy condition, 13 out of the 16 children correctly identified Grover as the incompetent puppet (81.25%; 95% CI: 62.5–100%); the remaining three children identified Cookie Monster as the incompetent puppet. See Fig. 2.

Children's ability to distinguish agents' competencies 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 use the information about differences in the agents' subjective rewards 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.

We asked the children "Which puppet cannot climb?" (rather than, for instance, "Which puppet has more difficulty climbing?") because what was at stake in this experiment was only children's ability to use the expected reward information to distinguish the expected costs for two puppets. Even the information about the differences in the agents' subjective rewards does not provide evidence about the agent's absolute competence. That is, the utility functions do not distinguish a puppet that is completely unable to climb the tall box from a puppet that

merely finds it very costly to do so. Critically however, children were able to recognize that the differences in agents' subjective preferences could provide information about subjective costs for the puppet who preferred one treat to another, but was unlikely to be informative for the puppet who liked both treats equally. This suggests that children understand how graded differences in subjective rewards can provide information about agents' subjective costs.

4. Experiment 3

Experiments 1 and 2 suggest that children are able to represent and infer agent-specific costs and rewards. In Experiment 3, we further investigate children's understanding of agent-independent (external) and agent-dependent (subjective) costs by asking whether children can manipulate the external cost associated with obtaining different goals to gain information about an agent's subjective costs. Children are given a high-reward and a low-reward treat and asked to place them in locations that incur different objective costs in order to learn an agent's subjective expected costs. Intuitively, it is obvious that agents should choose low cost actions when they are associated with high reward; it is therefore more informative to see how agents behave in the context of high reward, high cost actions. If children understand how costs and rewards affect an agent's choices, they should pair the high-reward treat with the high-cost location.

4.1. Methods

4.1.1. Participants

Sixteen 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 ($N = 8$).[2] One additional child was tested but excluded from analysis and replaced because she did not follow the instruction to place one item in each location (see Sections 2.2, 3.2, 4.2 and 5.2).

4.1.2. Stimuli

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

4.1.3. Procedure

Participants were tested in a quiet room at the museum in the presence of their caregiver. 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 (or Grover) does and see if he can climb. Where do you want to put them?"

4.2. Results and discussion

All videotapes were coded blind to condition by the first author for inclusion and children's responses to the test question; 81% of the videotapes were recoded on both measures by a second coder blind to hypotheses and conditions. The parents of one child did not consent to videotape and the child's response was judged online. Intercoder agreement was 100%.

As predicted, 14 of the 16 children made the informative intervention, putting the object with higher subjective reward in the more costly position (87.5% 95% CI: 75.0–100%). The remaining two children made an uninformative intervention, placing the object with the lower subjective reward in the more costly position. See Fig. 2. This suggests that children can predict how agents might act in the world as a function of the costs and rewards and can 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 had to combine a high-reward (HR) and a low-reward (LR) object with a high-cost (HC) and a low-cost (LC) location to generate a utility function. Although climbing the tall box is always more costly than climbing the short box ($HC > LC$), the exact difference between these costs is unobservable and variable across agents. For agents with high competence, this cost difference ($HC - LC$) is small. However, the less competent an agent is, the higher this cost difference becomes. If children place the high-reward object on the low-cost location, the agent can choose between a high-reward for a low-cost plan ($HR - LC$), and a low-reward for a high-cost plan ($LR - HC$). Here the agent's competence plays no role; it is always better to choose the high-reward for a low-cost plan (because $HR - LC > LR - HC$ for all values since $HR > LR$ and $HC > LC$). Thus the choice between these two plans reveals nothing about the agent's competence.

If, instead, children place the high-reward object at the high-cost location, then the agent's rational action choice becomes dependent on his competence. If the agent is very competent, then the high-reward for a high high-cost plan is likely to have a higher utility than the low-reward for a low-cost plan ($HR - HC > LR - LC$). However, if the agent is less competent, then the difference between the high-cost plan and the low-cost plan is relatively large ($HC - LC$) and the low-reward for a 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. Again, note that even the informative intervention does not provide evidence about the agent's absolute competence. As in Experiment 2, the agent might be unable to climb the tall box or merely find

it very costly to do so. Nonetheless, children were able to distinguish the more and less informative intervention and use information about agent's subjective rewards to provide evidence about the agent's competence.

5. Experiment 4

In Experiment 3, children manipulated the objective costs associated with each reward to make inferences about the agent's subjective costs. In Experiment 4 we hold the objective cost associated with each reward constant and ask instead whether children can identify agents whose subjective rewards are informative about their subjective costs. Following the same logic described above, if an agent assigns the same reward to two objects, that agent's actions are unlikely to be informative about his or her subjective costs: provided the agent is maximizing utilities, he will always choose the reward associated with lower cost. However, if an agent assigns a higher reward to one object than another, then if the agent fails to pursue the higher reward object, it suggests that the expected cost of the action was high.

In Experiment 4, children are shown two treats, one in a high cost location and one in a low cost location. Children are introduced to two puppets with different preferences and asked to identify the puppet who can perform the high cost action. If children can predict how an agent will act as a function of the costs and rewards, they should select the puppet that prefers the treat on the high-cost location. 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 have each treat be the favorite of one of the puppets.

5.1. Methods

5.1.1. Participants

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

5.1.2. Stimuli

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

5.1.3. Procedure

Participants were tested in a quiet room at the museum in the presence of their caregiver. 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?"

5.2. Results and discussion

All videotapes were coded by the first author blind to condition for inclusion and children's responses to the test question; 100% of the videotapes were recoded by a second coder on both measures blind to hypotheses and conditions. All parents consented to videotape. Intercoder agreement was 100%.

In Experiment 4, 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 (75%; 95% CI: 56.25–100%); the remaining four made the uninformative intervention (choosing the other puppet). See Fig. 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 was more likely to 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 more likely to choose the less preferred treat on the short box. Additionally, in this setup each treat was preferred by one of the agents. As such, children could not have succeeded through simpler strategies like associating the high-reward treat with a location out of reach, or ignoring the low-reward treat (as the low-reward treat was dependent on the agent). 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. Again, we asked about the puppet's ability to climb the tall box, because we were primarily interested in whether children could use the information about the agent's preferences to distinguish the two puppets. We do not know whether children interpreted the high cost action merely as very difficult for the puppet or as so costly that the agent was unable to perform the action at all. Critically however, children recognized that the intervention on the puppet that preferred the high cost treat was potentially informative whereas the intervention on the puppet that preferred the low cost treat was not.

6. 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 to maximize overall utilities and not just rewards, and as a consequence, 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.

In all experiments, children's success rate was consistently above chance (as assessed by 95% CIs). However, in several cases, the 95% confidence intervals were close to 0.5, raising a concern about the robustness of these effects. To test the overall evidence the four experiments provided for our theory, we conducted a Bayesian meta-analysis of effect sizes across all experiments using a hierarchical random-effects model. The results placed extremely high confidence on an estimate of children's overall rate of theory-correct responding being well above chance, with a mean posterior estimate of 78%, and 95% posterior CIs between 69% and 87%.[3] The variability in children's responses might be due to some but not all of the children having a mature naïve utility calculus, or due to other factors influencing children's responses in our tasks, as in any complex real-world judgment. However, taken together, the experiments support the hypothesis that children's judgments of agents' preferences and costs are consistent with intuitive utility calculations.

Our proposal of a naïve utility calculus is a natural extension of current accounts of goal-directed action understanding. As discussed, 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; Perner & Roessler, 2012). However, such early social

---

[3] In this analysis we assumed that each experiment followed a binomial distribution with unknown bias θ. However, rather than assuming that each experiment's bias was independent, we assumed that these biases were distributed in accordance with a beta distribution with parameters $\alpha = \mu k$ and $\beta = (1 - \mu)k$. Under this parameterization (following Kruschke, 2010) the beta distribution has mean value μ. Thus, if our data is evidence that children succeed in tasks requiring them to reason about the costs and rewards underlying rational action, then μ should be reliably greater than 0.5. As k, increases, μ quickly converges to a value of 0.78 with 95% CI: 0.69–0.87, suggesting that altogether, our four experiments provide strong evidence toward our theory. See Supplemental materials.

cognition has generally been demonstrated in the context of agent-independent, directly observable, situational and behavioral constraints. Even such sophisticated findings as early false belief understanding are predicated on knowing, for instance, that one person can see the contents of a box and one cannot (Onishi & Baillargeon, 2005; Southgate et al., 2007, 2010). In contrast, a mature naïve utility calculus requires understanding that even given identical epistemic access and situational contexts, agents differ in their subjective rewards and costs in ways that may affect their behavior.

Consistent with other research showing the emerging understanding of individual differences (Berndt & Heller, 1986; Kalish, 2002; Liu et al., 2007; Rholes & Ruble, 1984; Rotenberg, 1980, 1982; Seiver et al., 2013), and the ability to impute different perspectives on identical evidence (Carpendale & Chandler, 1996; Chandler & Helm, 1984), the current findings suggest that five and six-year-olds are sensitive to the internal structure of goals and recognize both agent-invariant and agent-dependent aspects of costs and rewards. Children always saw agents who were fully informed about the location of desired objects, they never saw agents change their minds, or attempt and fail to execute an action and agents always reached their goals successfully. Nevertheless, children were able to impute different unobservable motivations and competencies to the agents. Such results suggest that at least by the age of six, the principle of rational action extends to a principle of rational choice. Children not only expect agents to take rational actions toward their goals, but also to use the expected costs and reward of actions to decide when it is worth pursuing a goal at all. Further research might look at the developmental trajectory of these abilities to see if aspects of the naïve utility calculus emerge even earlier in development.

These studies also raise several questions for further research. First, our studies suggest that children understand that costs and rewards underlie goal formation, and that these costs and rewards vary across agents. However, the precise computations underlying these inferences are still unknown. Our account is motivated by Bayesian inference formalizations of goal-directed action understanding. However, scenarios with more complex and graded levels of preference and competence are needed to test the quantitative predictions of a Bayesian account. Moreover, the Bayesian account is a computational level analysis (Marr, 1982); the precise algorithms by which children represent and integrate information about costs and rewards remains a topic for future investigation. Second, although our experiments suggest that children can infer either subjective rewards (Experiment 1) or costs (Experiments 2–4) when the other factor is fixed, we do not know whether, given richer situational constraints and sequences of observable actions, children might also be able to do joint inference and simultaneously infer the costs and rewards that would explain the agent's actions.

Finally, although our results suggest that children understand that both costs and rewards vary across agents, we do not know if they understand that some aspects of the costs and rewards are more stable than others.

Agent-specific aspects of the costs and rewards can include both state- and trait-like differences. An agent might assign a high-expected cost to an action because of a transient state change (e.g., twisting an ankle) or because of a more stable trait (e.g., being weak or lazy). Similarly, some rewards have high value at some moments but not others (e.g., food when hungry) whereas other rewards may be more stable across time (e.g., long-term values or preferences). Further research might look at the development of children's understanding of both transient and stable aspects of subjective costs and rewards.

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 subjective costs and rewards of agents' actions, but can also act on the world gain information about these unobservable variables. These abilities emerge relatively early in development, in the absence of formal instruction, suggesting that a naïve utility calculus may be a fundamental component of human social cognition.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cognition.2015.03.006.

References

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).

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

Bellman, R. (1957). *A Markovian decision process* (No. P-1066). Santa Monica, CA: RAND CORP.

Berndt, T. J., & Heller, K. A. (1986). Gender stereotypes and social inferences: A developmental study. *Journal of Personality and Social Psychology, 50*(5), 889.

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.

Carpendale, J. I., & Chandler, M. J. (1996). On the distinction between false belief understanding and subscribing to an interpretive theory of mind. *Child Development, 67*(4), 1686–1706.

Chandler, M. J., & Helm, D. (1984). Developmental changes in the contribution of shared experience to social role-taking competence. *International Journal of Behavioral Development, 7*(2), 145–156.

Cohen, J. (1994). The earth is round (p < . 05). *American psychologist, 49*(12), 997.

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.

Csibra, G., & Gergely, G. (1998). The teleological origins of mentalistic action explanations: A developmental hypothesis. *Developmental Science, 1*(2), 255–259.

Cumming, G. (2013). The new statistics why and how. *Psychological Science*. 0956797613504966.

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

Gergely, G., & Csibra, G. (1997). Teleological reasoning in infancy: The infant's naive theory of rational action: A reply to Premack and Premack. *Cognition, 63*(2), 227–233.

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naive 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*.

Kalish, C. W. (2002). Children's predictions of consistency in people's actions. *Cognition, 84*(3), 237–265.

Kruschke, J. (2010). *Doing Bayesian data analysis: A tutorial introduction with R*. Academic Press.

Liu, D., Gelman, S. A., & Wellman, H. M. (2007). Components of young children's trait understanding: Behavior-to-trait inferences and trait-to-behavior predictions. *Child Development, 78*(5), 1543–1558.

Marr, D. (1982). Vision, 1982. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information.*

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308*(5719), 255–258.

Perner, J., & Roessler, J. (2012). From infants' to children's appreciation of belief. *Trends in Cognitive Sciences, 16*(10), 519–525.

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

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: evidence from 14-and 18-month-olds. *Developmental Psychology, 33*(1), 12.

Rholes, W. S., & Ruble, D. N. (1984). Children's understanding of dispositional characteristics of others. *Child Development*, 550–560.

Rotenberg, K. J. (1980). Children's use of intentionality in judgments of character and disposition. *Child Development*, 282–284.

Rotenberg, K. J. (1982). Development of character constancy of self and other. *Child Development*, 505–515.

Ruble, D. N., & Dweck, C. S. (1995). Self-conceptions, person conceptions, and their development. *Social Development*, 15.

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.

Seiver, E., Gopnik, A., & Goodman, N. D. (2013). Did she jump because she was the big sister or because the trampoline was safe? Causal inference and the development of social attribution. *Child Development, 84*(2), 443–454.

Skerry, A. E., Carey, S. E., & Spelke, E. S. (2013). First-person action experience reveals sensitivity to action efficiency in prereaching infants. *Proceedings of the National Academy of Sciences, 110*(46), 18728–18733.

Sommerville, J. A., Woodward, A. L., & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition, 96*(1), B1–B11.

Southgate, V., Chevallier, C., & Csibra, G. (2010). Seventeen-month-olds appeal to false beliefs to interpret others' referential communication. *Developmental Science, 13*(6), 907–912.

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science, 18*(7), 587–592.

Tomasello, M., Carpenter, M., Call, J., Behne, T., Moll, H., et al. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences, 28*(5), 675–690.