What's worth the effort: Ten-month-old infants infer the value of goals from the costs of actions

Shari Liu[1] (shariliu01@g.harvard.edu), Tomer D. Ullman[1,2] (tomeru@mit.edu), Joshua B. Tenenbaum[2] (jbt@mit.edu), & Elizabeth S. Spelke[1] (spelke@wjh.harvard.edu)
[1]Department of Psychology, Harvard University, Cambridge, MA 02138
[2]Department of Brain and Cognitive Sciences, MIT, Cambridge, MA 02139

Abstract

Infants understand that people act in order to achieve their goals, but how can they tell what goals people find worthwhile? Here, we explore the thesis that human infants solve this problem by building a mental model of action planning, taking into account the costs of acting and the rewards actions bring. Consistent with this thesis, we found that 10-month-old infants, after viewing an agent approach two objects equally often, inferred that the agent preferred the object whose attainment required a costlier action. Infants' responses generalized across changes in perceptual variables that distinguished one action from another (e.g. path length, angle of incline), suggesting that an abstract cost metric based on *force* or *effort* supported their judgments. These findings suggest that infants' knowledge about agents may be expressed as a generative model for action planning, which can then be inverted to identify the probable hidden causes for observed actions.

Keywords: social cognition; cognitive development

Introduction

When we observe the actions of others, we see more than bodies moving in space. A hand reaching for an apple is not just one object decreasing its distance from another object; it can indicate hunger (in the person who is reaching), helpfulness (if the person is reaching on behalf of someone else) or compromise (if the person reaching would prefer a banana, but not enough to go buy one). This fast, automatic and consistent ability to interpret the behavior of others as intentional, goal-directed and constrained by the physical environment is often dubbed 'intuitive psychology' (e.g. Baker, Saxe, & Tenenbaum, 2009; Dennett, 1987; Gergely & Csibra, 2003; Heider & Simmel, 1944). It raises questions about the developmental origins and computational basis for these judgments: How does this faculty develop and change over time? How can it be formalized? And what algorithms support its efficient use?

Over the past two decades, research has revealed that the building blocks of our intuitive psychology are present in the first year of life. Despite infants' limited experience observing and acting on the world, their understanding and learning is guided by assumptions about the physical properties (Saxe, Tzelnic, & Carey, 2006), intentions and goals (Woodward, 1998), mental states (Luo & Johnson, 2009; Onishi & Baillargeon, 2005), and causal powers (Muentener & Carey, 2010) of agents, as well as the cost associated with their actions (Gergely & Csibra, 2003; Liu & Spelke, 2017). This wealth of findings raises a fundamental question about early intuitive psychology: Do infants' capacities testify to distinct local abilities (Heyes & Frith, 2014; Paulus, 2012; Woodward, 2009), or to a single coherent system supporting inference, prediction, and learning (Baillargeon, Scott, & Bian, 2016; Carey, 2009; Gergely & Csibra, 2003; Spelke & Kinzler, 2007)?

Here, we tackle this question in a case study, based on a computationally precise proposal for a coherent, abstract, and productive system for action understanding (Fig. 1). We ask whether infants, like adults, infer the hidden causes of agents' actions by assuming that agents make plans to maximize rewards and minimize costs (Baker et al., 2009; Dennett, 1987; Jara-Ettinger, Gweon, Schulz, & Tenenbaum, 2016). While infants have been shown to understand the preferences of agents (Woodward, 1998) and the efficiency of their actions (Gergely & Csibra, 2003; Liu & Spelke, 2017), it remains open whether their understanding reflects a unified, abstract system that reasons over variables like effort and value, or piecemeal assumptions about simple cues like distance and time.

We present two experiments that investigate the content and productivity of infants' intuitive psychology. Under the framework of the naïve utility calculus (Jara-Ettinger et al., 2016), rational agents tend to act only when the utility of these actions is greater than zero: when the rewards that their actions bring outweigh their costs. Thus, the costs agents incur to transition to certain states of the world provide a lower limit for the value of these states: when a person chooses to walk 10 feet, 1 mile, or 5 miles to the coffee shop, we can infer that her desire for coffee is at least as strong as the energy required for her to make the trip. This suggests that one of the most reliable signals of relative subjective value—how much or how little an agent likes a certain state of the world—is the amount of effort the agent is willing to expend in order to bring it about or prevent it. In the present research, we ask whether infants (a) infer the relative value of two goals from the relative effort agents expend in order to achieve them and (b) use this inferred information to generate a prediction about which goal the agent will choose under conditions of equal cost.

Experiment 1

Here, we ask whether infants use the relative costs that an agent incurs to achieve a goal to infer the relative value of that goal to the agent. We showed infants that an agent was willing to pay a higher cost to reach one goal compared to another. We then measured infants' attention when they saw the agent choosing the higher value goal over the lower value

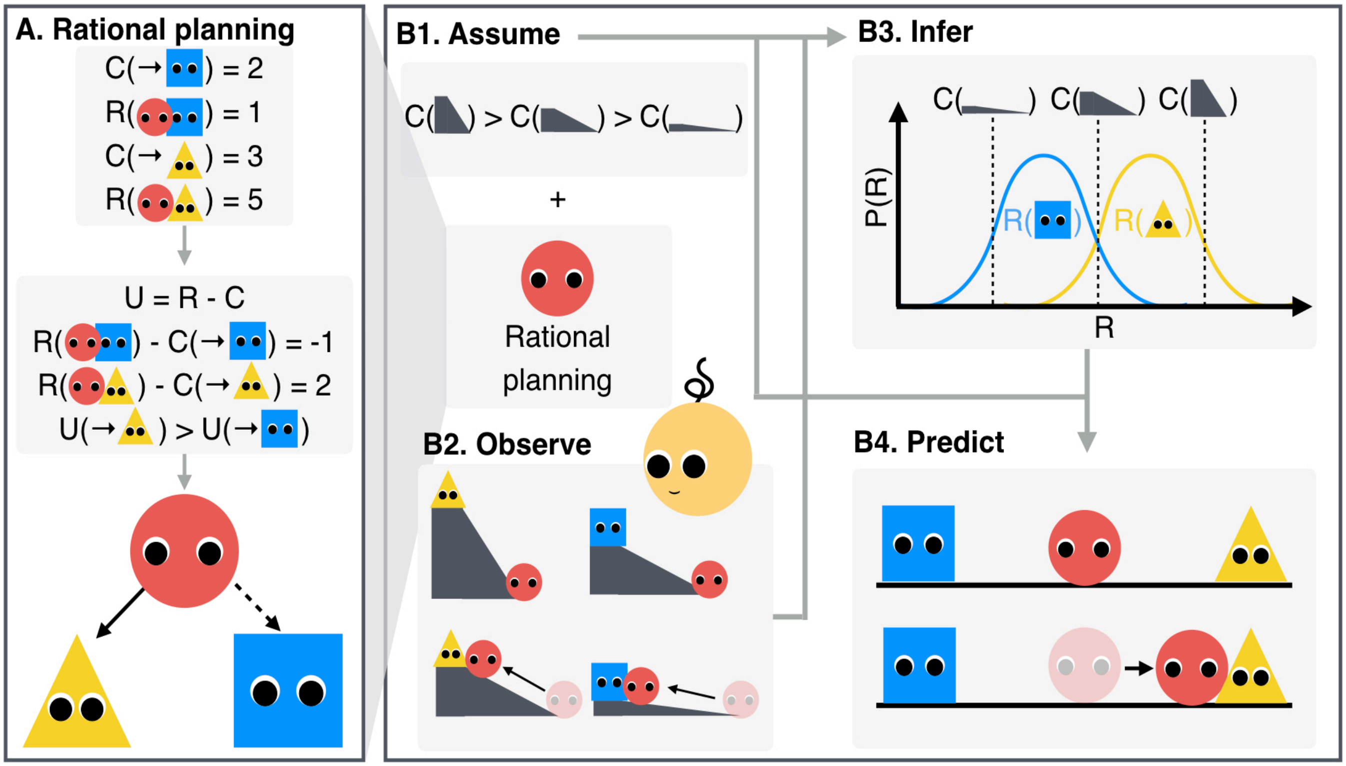

Figure 1: A schematic of our computational framework. The forward direction (A) defines the agent as a rational planner that calculates the utilities of different actions from their respective costs and rewards, and then selects an action stochastically in proportion to its utility. In this case, the overall utility for approaching Triangle is higher than for approaching Square, so the Protagonist likely will choose Triangle over Square. In the inverse direction (B1), an observer assuming this model and some priors over the costs of different actions, can (B2) observe a series of actions and then (B3) infer a posterior distribution over the hidden values of an agent's costs and rewards given its actions. These posteriors can then be used to (B4) predict the actions of the agent in a new situation, in which the same goal states can be reached by different actions.

goal, and vice versa. We predicted that if infants can infer value from effort and use this information to predict the agents' subsequent actions, then they will differentiate between the more probable outcome (when the agent chooses the higher-value goal) and the less probable one (when the agent chooses the lower-value goal) by showing a looking preference in either direction.

Methods

Participants Our final sample included 24 healthy, full-term infants (15 female, M_{age}=9.95 months, range=9.43-10.53). Eight more infants were tested, excluded, and replaced (1 for fussiness that prevented study completion, 1 for technical failure, 5 for inattentiveness during test events, and 1 for experimental error). Sample size and exclusion criteria were fixed prior to the start of data collection, and decisions concerning exclusions were made by researchers unaware of the order of the test events viewed by the infant (and therefore blind to the data that the infant provided). All participants were recruited from the greater Boston area and tested at the Laboratory for Developmental Studies at Harvard University with parental informed consent. Families received a small thank-you gift (e.g. a t-shirt or toy) for participating.

Materials and Design All animated events were created in Blender (Blender Foundation, 2014), synchronized with a custom audio track in iMovie, and presented using Keynote on a 40" by 52" LCD projector screen. Two speakers flanking the screen played all stimuli-related sounds. Infants' looking time data were coded online using Xhab64 (Pinto, 1995), and offline using jHab (Casstevens, 2007).

The experiment consisted of 3 pairs of familiarization trials, 1 pre-test trial, and 2 pairs of test trials. All trials began with an attention-getting animation and sound (3.0s), followed by looped sequences of events. *Familiarization* sequences consisted of 4 videos (8s or 8.9s each; see below) and *test* sequences consisted of a single video (5.2s). Black screens (0.5s) were interspersed between events. Events featured three agents: a central red agent (the Protagonist), and two target agents (Triangle and Square). The left-right locations of the targets were constant across participants, but the identity of the higher-value target was counterbalanced across participants.

During *familiarization* (Fig. 2A), the Protagonist responded to the call of one of the targets by accepting or refusing to jump over a small (1 units tall), medium (6 units), or large (10 units) barrier that fell with a thud between the Protagonist and target. When the Protagonist accepted the cost (8.9s), it looked up at the barrier, made a positive "Mmm!" sound, and leapt over it to reach the target. The Protagonist always acted efficiently, adapting the height of its jump to the height of the barrier; all jumps were accompanied with a popping sound. When the Protagonist refused the cost (8.0s), it looked up at the barrier, made a mildly negative "Hmm…" sound, and backed away, returning to the center of the screen. During each familiarization trial, the Protagonist accepted a small cost and refused a medium cost for one target, and accepted a medium cost and refused a large cost for the other target. Each familiarization trial consisted of sequences presented in the above order (small, medium, medium, large) or the opposite order, and across 6 familiarization trials, both orders were presented 3 times in an ABABAB pattern. The identity of the higher-value target and the first familiarization sequence were counterbalanced across participants. Thus, the Protagonist approached and refused each target equally often, with equal affect, but took a costlier action for one of the targets.

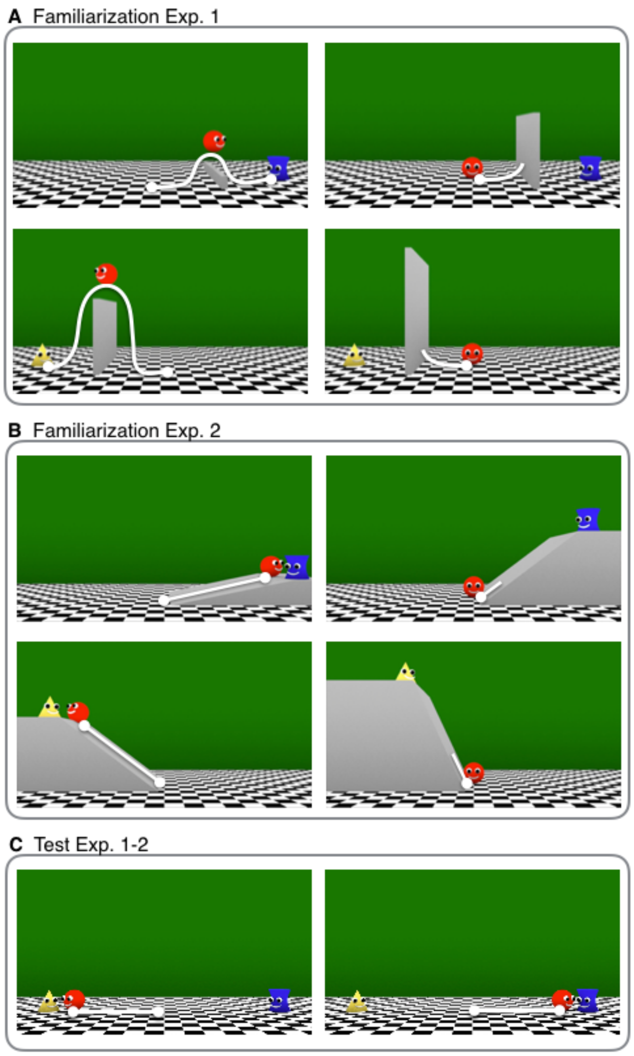

Figure 2: Structure of Experiments 1-2. During familiarization (A-B), the Protagonist (Circle) accepted a low and refused a medium cost for the lower-value target (in this case, Square), and accepted a medium and refused a high cost for the higher-value target (Triangle). At test (C), the Protagonist looked at each of the two targets and chose either the lower or higher value target. White circles indicate start- and end-points of action, and white lines indicate trajectories.

A *pre-test* event following familiarization featured a still image of the two targets without the Protagonist. During *test* (Fig. 2C), the Protagonist reappeared, rotated left then right while saying "Hmmm…", and then approached one of the targets: either the higher- or lower-value target. The same sound accompanied each approach. Across 2 pairs of test trials presented in ABAB order, the Protagonist approached the yellow Triangle target twice and the blue Square target twice. The first test trial (higher- or lower-value choice) was counterbalanced across participants.

Procedure Infants were seated on their caregivers' laps approximately 60" away from the screen. Caregivers were instructed to keep their eyes closed and to refrain from interacting with their infants throughout the experiment, and were monitored for compliance.

The researcher ran the experiment and coded looking time online while unaware of the order of events (and therefore unaware of the infant's differential reactions to the displays), but could determine the start of each trial as well as the general timing of actions (e.g. when the Protagonist approached one of the two targets), but no information about specific actions (e.g. which target the Protagonist chose), based on auditory cues.

Across both the familiarization and test phases of the experiment, the researcher began coding a trial immediately following the attention getter, and concluded the trial once the infant had attended to the screen for 60 cumulative seconds or looked away for 2 consecutive seconds. During pre-test, the researcher waited until the infant looked towards each target agent at least once, and then began the test trials. These criteria were fixed prior to the start of data collection.

Coding and Analysis Condition-blind observers coded videos from test sessions offline, and reviewed them for predetermined exclusion criteria (fussiness that prevented study completion, online coding error, experimenter error, technical failure, and parental interference). If infants were determined to have missed a critical, predetermined part of the test trial, then that test pair was marked and excluded from subsequent analyses. If all test trials were excluded, infants were replaced. All test trials were recoded independently by an additional researcher who was unaware of test pair order. The two coders agreed on trial cutoffs for 94.48% of the test trials, and the intraclass correlation (ICC) between the two raters was 0.994, [0.991, 0.996]. Thus, the primary offline coding data were used in our analyses.

Our primary dependent measure, specified before data collection began, was log-transformed looking time (Csibra, Hernik, Mascaro, Tatone, & Lengyel, 2016) but plots and descriptive statistics feature raw values for ease of interpretation. To explicitly take into account repeated measures, all linear mixed effects models (Bates, Mächler, Bolker, & Walker, 2014) included participants as a random intercept. Three classes of models were fit: (1) null models, featuring participant identity as the only predictor, (2) hypothesis-driven models, which included theory-relevant manipulated factors, and (3) exploratory models, which included additional non-hypothesis driven predictors. We leveraged likelihood ratio tests (LRTs) to evaluate model fit by assessing whether the inclusion of certain predictors significantly reduced residual variance. All model-produced degrees of freedom were calculated using the Satterthwaite approximation method. All reported p-values in Exp. 1 are two-tailed. Bracketed values indicate 95% confidence intervals.

Results

Hypothesis-driven Results A model with the single predictor of test trial (higher- or lower-value) revealed that infants looked longer at the lower-value action (M=28.41s, SD=12.29) than the higher-value action (M=21.79s, SD=14.85), B=0.327, SE=0.130, β=0.502, $t(24)$=2.523, p=.019, [0.062, 0.591]. This model outperformed a null model by a LRT, $X^2(1)$=5.648, p=.017. A leverage analysis using Cook's distance (Nieuwenhuis, te Grotenhuis, & Pelzer, 2012) revealed 1 influential observation in this model. Removal of this case produced an inferentially equivalent result, B=0.263, SE=0.119, β=0.454, $t(23)$=2.221, p=.037, [0.021, 0.506].

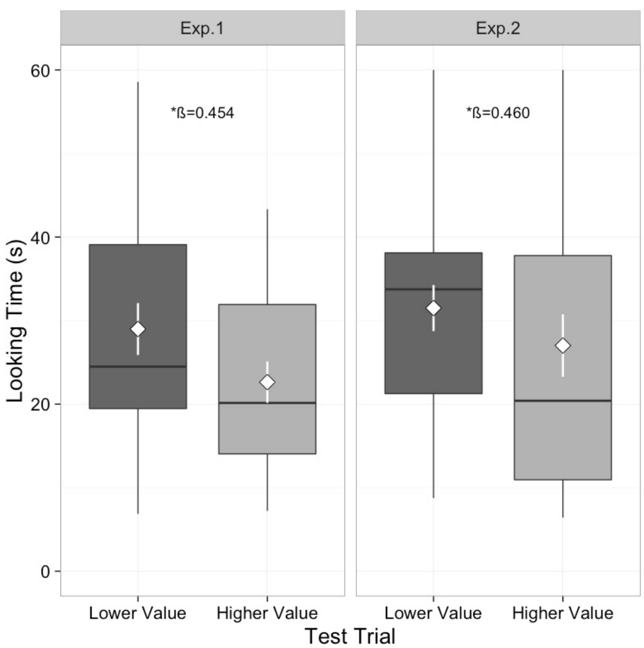

Figure 3: Boxplots of average looking times in seconds towards the higher and lower value test events in Experiments 1 and 2. Boxes indicate middle quartiles, and vertical lines indicate full range of values, and horizontal lines indicate medians. Means and standard errors are plotted in white. Beta coefficients indicate effect sizes in standard deviation units, and asterisks indicate two-tailed p-values (*<.05).

Exploratory Results A first exploratory analysis tested for an effect of test pair order by including an interactive effect between test trial presentation order and trial type. Infants discriminated between the test events to a similar degree regardless of whether they were assigned to watch the lower- or higher-value event first, B=0.212, SE=0.255, β=0.325, $t(24)$=0.829, p=0.415, [-0.310, 0.733]. Removing one influential case produced an inferentially equivalent result, B=0.354, SE=0.226, β=0.610, $t(23)$=1.569, p=0.130, [-0.107, 0.816]. A second model was fit with the additive effects of test pair order and test trial type, summed looking time during familiarization, sex, the identity of the higher value character, and the order of the first block of familiarization. Infants' looking preferences were not predicted by any of the exploratory factors, with all CIs containing 0, ps>0.1. Removal of one influential case produced inferentially equivalent results. The best model out of the above 4 was the hypothesis-driven model (AIC=92.500)

Discussion

In Exp. 1, we found that infants looked longer when the agent moved to the lower-valued target, providing evidence that they inferred the relative value of two different goals of an agent and used these inferred variables to predict its subsequent actions. This result is consistent with the thesis that infants represent the costs and rewards of actions under a single utility function, but leaves open two questions. First, did the infants respond to the curvature or length of the agent's trajectory rather than its cost (though many previous experiments cast doubt on this possibility, e.g. Gergely et al., 1995; Liu & Spelke, 2017; Skerry, Carey, & Spelke, 2013)? Second, and more importantly, do infants represent the cost of an action in terms of a simple feature, like path length, or in more abstract terms, like the force required to generate actions? Experiment 2 addressed both of these questions by providing a conceptual replication of Exp. 1.

Experiment 2

Exp. 2 was identical to Exp. 1 in materials, design, and procedure except that the agent accepted and refused to travel up ramps to reach its goal when the ramps varied in incline but were equal in length. If infants compute the cost of the agent's actions rather than more superficial features of the agent's motion such as its speed, path curvature, or path length, then they should perform in this experiment as they did in Exp. 1.

Methods

Participants Our final sample included 24 healthy, full-term infants (15 female, M_{age}=9.88 months, range=9.47-10.43). Ten more infants were tested, excluded, and replaced (1 for fussiness that prevented study completion, 1 for technical failure, 2 for online coding errors, 2 for parental interference, and 4 for inattentiveness during test events).

Materials, Design, and Procedure During *familiarization* (Fig. 2B), each target appeared at the top of a ramp and the Protagonist either climbed or refused to climb the ramp to reach it. To manipulate cost while controlling for path length, the angle of the ramps was varied (11.51°, 39.26°, and 64.09°) such that the inclined plane was always 10 Blender units in length. When the agent accepted the cost (6.8s), it moved up the ramp once, slid back down, and then moved to the target. When the agent refused the cost (5.5s), it moved up the ramp once, slid back down, and then turned away from the target. Thus, the agent approached each object equally often by moving equally far.

Coding and Analysis All test events were recoded as for Exp. 1. The two coders agreed on trial cutoffs for 97.89% of the test trials; the ICC between the two raters was 0.978, [0.967, 0.985]. Thus, the primary offline coding data were used in our analyses. Because of our strong directional prediction, all reported *p*-values in hypothesis-driven results in Exp. 2 are one-tailed. All other reported *p*-values are two-tailed.

Results

Hypothesis-driven Results As in Exp. 1, infants looked longer at the lower-value action (M=30.51s, SD=13.79) than the higher-value action (M=27.05s, SD=17.55), B=0.221, SE=0.121, β=0.350, $t(24)$=1.826, p=.040, [-0.026, 0.468]. This model outperformed a null model by LRT, $X^2(1)$=3.121, p=.077. Removal of 1 influential case produced an inferentially equivalent result, B=0.286, SE=0.108, β=0.460, $t(23)$=2.660, p=.007, [0.066, 0.506].

Exploratory Results An exploratory model testing explicitly for presentation order revealed that infants differentiated between the test events differently depending on whether they were assigned to watch the lower-value versus the higher-value approach first, B=0.579, SE=0.211, β=0.917, $t(24)$=2.739, p=.011, [0.147, 1.011]. We detected 1 influential observation in this model and removed it from subsequent pairwise comparisons, which revealed that whereas infants who saw the low value test event first looked longer at the low value (M=28.70s, SD=11.12) versus high value (M=20.01s, SD=14.83) test trials, B=0.511, SE=0.139, $t(25.19)$=3.679, p=.001, [0.225, 0.796], infants who saw the high value choice did not differentiate between the lower (M=34.56s, SD=15.09) and higher-value test events (M=34.65, SD=18.53), B=0.041, SE=0.145, $t(25.19)$=0.283, p=0.779, [-0.257, 0.339]. An additional model testing for effects of summed attention during test, sex, the identity of the higher value target, and the first familiarization loop revealed that no further effects other than one of first familiarization loop, where infants assigned to watch a sequence of low to high cost first looked marginally longer overall at test, B=0.392, SE=0.194, β=0.520, $t(24)$=2.021, p=.055, [0.147, 1.010]. The best model of the four reported was the simpler exploratory model with the single interactive effect (AIC=84.648).

Comparing Exp. 1 and 2 Across both experiments, infants looked longer at the lower value action (M=29.46s, SE=14.22) than the higher value action (M=24.42s, SD=15.22), B=0.274, SE=0.089, β=0.427, $t(48)$=3.079, p=.003, [0.096, 0.452]. Removal of one influential case produced an inferentially equivalent result, B=0.242, SE=0.085, β=0.400, $t(47)$=2.849, p=.006, [0.072, 0.412]. An additional model with an interactive effect between experiment and test event was fit and revealed no differences in looking preference across experiments, B=0.106, SE=0.1773, β=0.165, $t(48)$=0.597, p=.553, [-0.249, 0.460]. Removal of one influential observation produced an inferentially equivalent result, B=0.042, SE=0.170, β=0.070, $t(47)$=0.250, p=.804, [-0.297, 0.382]. Regardless of whether infants saw the central agent jump higher barriers (Exp. 1) or climb steeper ramps (Exp. 2) for one target over another, they expected the agent to approach the higher-value target at test.

General Discussion

Across two experiments, infants reasoned jointly about effort and value. We found that 10-month-old infants successfully derived the relative value of two potential goals from evidence that an agent was willing to take a higher cost for one goal than the other, and used it to predict an agent's choice between them. Infants succeeded whether they were provided with evidence about the length (Exp. 1) or the steepness (Exp. 2) of the path the agent was willing to take to obtain the object, suggesting that rich variables like cost and value, rather than features like path length or action duration, support their judgments.

These experiments advance our understanding of core cognition in several ways. First, they support the thesis that an abstract, coherent and productive system guides the analysis of agents and their actions (Baillargeon et al., 2016; Carey, 2009; Gergely & Csibra, 2003), and go beyond them by providing and testing a proposal for the specific form and content that express the productivity and abstraction in these systems. First, while preverbal infants are known to represent the preferences of agents (Woodward, 1998) and the efficiency of their actions (Gergely & Csibra, 2003), we provide the first evidence that they represent the utilities of actions that subsumes both costs and rewards (Jara-Ettinger et al., 2016) and assume that agents plan over these variables (Baker et al., 2009). Second, in accord with evidence that infants represent physical cost as a continuous variable (Liu

& Spelke, 2017), we found that rewards are represented at least ordinally. Next, our findings suggest that infants represent physical cost in terms of the forces that cause the actions of agents, rather than the length or duration of the trajectory that these actions follow. Experiments manipulating other aspects of cost (e.g. mental effort, physical risk) and reward (e.g. number, transitivity) can shed further light on the content supporting infants' representations of effort and value. Taken together, our findings suggest that the cognitive machinery required to account for infants' performance requires a mental model both of how agents work in the forward direction (in accord with a principle of maximizing utilities; Jara-Ettinger et al., 2016) and a procedure for inverting this model (in accord with the computational framework of inverse planning; Baker et al., 2009). We hope in future work to develop and test such a model.

More generally, the present research suggests that computational frameworks provide a powerful guide for studying cognitive development, both in motivating and interpreting answers to central questions about the origins of knowledge. Using this approach, we demonstrated that far before human infants learn to walk, leap, and climb, they leverage rich mental models of utility to understand the actions of others, including notions of cost grounded in intuitive physics, and notions of reward that can be inferred from them.

Acknowledgements

This material is based upon work supported by the Center for Brains, Minds and Machines, funded by National Science Foundation STC award CCF-1231216, and by a National Science Foundation Graduate Research Fellowship under Grant No. DGE-1144152. We thank the families who volunteered to participate, A. Aguirre and C. Kerwin for research assistance, R. Guzman and N. Kalra for administrative assistance.

References

Baillargeon, R., Scott, R. M., & Bian, L. (2016). Psychological Reasoning in Infancy. *Annual Review of Psychology*, *67*(1), 159–186.

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

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1).

Carey, S. (2009). *The origin of concepts.* New York, NY: Oxford University Press.

Casstevens, R. M. (2007). jHab: Java Habituation Software (Version 1.0.0) [Computer software]. Chevy Chase, MD.

Csibra, G., Hernik, M., Mascaro, O., Tatone, D., & Lengyel, M. (2016). Statistical treatment of looking-time data. *Developmental Psychology*, *52*(4), 521-536.

Dennett, D. C. (1987). *The Intentional Stance*. London: The MIT Press.

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.

Heider, F., & Simmel, M. (1944). An experimental study of social behavior. *The American Journal of Psychology*, *57*(2), 243–259.

Heyes, C. M., & Frith, C. D. (2014). The cultural evolution of mind reading. *Science, 344*(6190), 1243091.

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

Liu, S., & Spelke, E. S. (2017). Six-month-old infants expect agents to minimize the cost of their actions. *Cognition*, *160*, 35–42.

Luo, Y., & Johnson, S. C. (2009). Recognizing the role of perception in action at 6 months. *Developmental Science*, *12*(1), 142–149.

Muentener, P., & Carey, S. (2010). Infants' causal representations of state change events. *Cognitive Psychology, 61*(2), 63-86.

Nieuwenhuis, R., te Grotenhuis, M., & Pelzer, B. (2012). Influence.ME: Tools for detecting influential data in mixed effects models. *R Journal, 4*(2), 38-47.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308*(5719), 255–8.

Paulus, M. (2012). Action mirroring and action understanding: an ideomotor and attentional account. *Psychological Research*, *76*(6), 760–7.

Pinto, J. (1995). Xhab64 [Computer software]. Palo Alto, CA: Stanford University.

Saxe, R., Tzelnic, T., & Carey, S. (2006). Five-month-old infants know humans are solid, like inanimate objects. *Cognition*, *101*(1), B1–B8.

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 of the United States of America*, *110*(46), 18728–33.

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

Stiching Blender Foundation (2016). Blender (Version 2.71) [Computer software]. Retrieved from https://www.blender.org/download/

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

Woodward, A. L. (2009). Infants' grasp of others' intentions. *Current Directions in Psychological Science*, *18*(1), 53–57.