# Ten-month-old infants infer the value of goals from the costs of actions

**Authors:** Shari Liu*[1,3], Tomer D. Ullman[1,2,3], Joshua B. Tenenbaum[2,3], & Elizabeth S. Spelke[1,3]

**Affiliations**:
[1]Department of Psychology, Harvard University
[2]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology
[3]Center for Brains, Minds, and Machines (CBMM)

*Correspondence to: shariliu01@g.harvard.edu

**Abstract:** Infants understand that people pursue goals, but how do they learn which goals people prefer? Here, we test whether infants solve this problem by inverting a mental model of action planning, trading off the costs of acting against the rewards actions bring. After seeing an agent attain two goals equally often at varying costs, infants expected the agent to prefer the goal it attained through costlier actions. These expectations held across three experiments conveying cost through different physical path features (jump height and width; incline angle), suggesting that an abstract variable, such as 'force', 'work' or 'effort', supported infants' inferences. We model infants' expectations as Bayesian inferences over utility-theoretic calculations, providing a bridge to recent quantitative accounts of action understanding in older children and adults.

**One Sentence Summary:** Infants use the amount of work or effort an agent is willing to expend, to determine the value of the agent's goal.

**Main Text**:

When we observe people's actions, we see more than bodies moving in space. A hand reaching for an apple is not just one object decreasing its distance from another; 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 and automatic ability to interpret the behavior of others as intentional, goal-directed, and constrained by the physical environment is often termed 'intuitive psychology' (*1–4*). Here, we use behavioral experiments and computational models to probe the developmental origins and nature of this ability.

Over the past two decades, research has revealed that the building blocks of our intuitive psychology are present as early as the first year of life. Despite infants' limited experience, their interpretations of other people's actions are guided by assumptions about agents' physical properties (*5*), intentions and goals (*6*), mental states (*7–10*), causal powers (*11*), and dispositions to act efficiently (*7*, *12*, *13*). This wealth of findings does not reveal, however, whether infants' capacities depend on a host of distinct local abilities (*14–16*), or on a single coherent system supporting inference, prediction, and learning (*3*, *17–19*).

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). Previous studies suggest that infants are sensitive to the costs of agents' actions (*3*, *7*, *12*, *13*), and can infer agents' preferences (*6*, *9*). Decision theorists for hundreds of years have recognized these as the two central factors guiding the decisions of rational agents (*20–22*). Here we ask whether infants can integrate these dimensions to infer agents' goals: Do infants use the cost that an agent expends to attain a goal state in order to infer the value of that goal state for the agent?

[Fig 1 here]

Such an inference has been proposed to rest on three nested assumptions that together constitute a "naïve utility calculus" (*23*), analogous to classical economic thinking. First, agents act to maximize their utility U, under constraints (*2*, *4*, *24*, *25*). Second, this utility separates into rewards and costs, two distinct components that can be individual targets of inference (*26*). That is, if R(S) is the reward of a goal state S, and C(A) is the cost of an action, then an agent acts to maximize the following:

$$U(A,S) = R(S) - C(A). \qquad (1)$$

Third, the cost of an action is not arbitrary but depends in general on properties of both the agent and the situation they are in, which determine how much effort the agent might need to exert to carry out that action.

These assumptions can be formalized as generative models that successfully predict the quantitative and qualitative behavior of adults and older children (*4*, *23*, *26*). In these models, observers who reason that other agents are maximizing their expected utility according to Eq. 1 can use what they know about rewards and costs to predict the agents' future actions. Inverting this process, observers can use the agents' overt actions to infer their hidden rewards and costs, according to:

$$P(R,C \mid A) \propto P(A \mid C,R) \cdot P(R,C), \qquad (2)$$

where $P(R,C|A)$ is the posterior distribution over the rewards and costs of an agent. By Bayes' theorem, this distribution is proportional to the product of $P(A|C,R)$ — the likelihood of the agent choosing action A given rewards R and costs C, given by a rational planning procedure (*4*, *23*) —and $P(R,C)$, a prior distribution over costs and rewards.

Do infants apply the logic of cost-reward reasoning? Past research suggests that infants are sensitive to the relative value of different goal objects for an agent who chooses to approach one object over another (*6*, *27*) as well as the relative efficiency of the actions taken by an agent who approaches a goal object (*12*, *13*, *28*). Past studies do not reveal, however, whether infants have a unified intuitive psychology in the form of a generative model, or separate representations for variables like cost and reward that become unified only later in development, as children gain experience exerting themselves to achieve goals or communicating with others about their desires and actions. It is also an open question whether infants consider cost and reward in terms of abstract variables, such as work, effort, desire or value*,* or whether their understanding is restricted to perceptual features of actions, such as the distance or duration an agent travels, or the number of times it selects a particular goal. In physical action contexts, effort often covaries with perceptible properties such as the length or duration of a path traveled, but it depends ultimately on the amount of force that the agent must exert over time and distance (i.e., the amount of work the agent must do). Likewise, value often covaries with the number of times agent selects a goal, but ultimately depends on how strongly the agent desires a goal relative to the cost of achieving it or its value relative to other options.

We designed and conducted three experiments to test whether infants learn about the reward agents place on goals from cost, working backwards from the assumption that agents maximize utility and inferring relative rewards from observed actions under varying costs. We then use the data from these experiments, together with the findings from past experiments (*6*, *7*, *13*), to test a variety of computational models of infants' performance, including models with integrated versus isolated, and abstract versus cue-based, representations of costs and rewards (see model description in Supplementary Materials). Our empirical and computational findings support the view that a productive system grounded in costs-reward tradeoffs guides action understanding toward the end of the first year of life.

[Fig 2 here]

We tested $N$ =80 ten-month-old infants in three experiments with pre-specified designs, procedures, sample sizes, and analysis plans (*29*). In all experiments, infants first saw an agent move to and refuse to move to each of two target goals under conditions of varying cost. Then infants watched test events in which the agent chose either the higher or the lower value target when both were present at equal cost. If infants infer the reward of the targets to the agent from the effort undertaken to reach them, and then they should be more surprised when the agent chooses the lower value target, looking longer at the test trials displaying that action (*30*).

In Experiment 1 ($N$=24), we leveraged events widely used in studies of early action understanding, wherein animated characters jump efficiently over barriers of variable heights to arrive at goal objects (*3*, *7*, *13*, *31*), and indicate their preferences by selecting one goal over another (*6*, *9*). During familiarization, infants watched six trials consisting of four different events involving a central agent and one of two target individuals on a level surface (Fig. 2A; Movie S1). In each event, the target jumped and made a noise, and the agent responded by turning to face and beginning to approach the target, whereupon a barrier fell onto the stage between directly in the agent's path. On two of these events (one for each target), the agent looked to the top of the barrier, made a positive "Mmmm!" sound, backed up and then jumped over the barrier, landing next to the target. On the other two events, the agent looked to the top of the barrier, made a neutral "Hmmm…" sound, backed away, and returned to its initial position. The critical distinction between these events concerned the height of the barrier, and therefore the length, height, and speed of the jump that the agent undertook so as to clear it (all jumps were equated for duration). For one target, the agent jumped over a low barrier and declined to jump a medium barrier; for the other target, the agent jumped the medium barrier and declined a tall barrier. After this familiarization, the agent appeared between the two equidistant targets on a level surface. Infants viewed two pairs of looped test events (Fig. 2D-2E, Movies S4-5), order counterbalanced, in which the agent looked at each of the targets and then repeatedly approached either the higher or the lower value target. Our pre-specified dependent measure was average log-transformed looking time (*32*) across test trials; we predicted differential looking at the test events but did not pre-specify the direction of this difference.

Infants looked longer at test trials in which the agent chose the target for whom it had jumped a lower barrier ($M$=28.41s, $SD$=14.85), relative to the target for whom it had jumped a higher barrier ($M$=21.79s, $SD$=12.29) (Fig. 3), 95% CI [0.062, 0.591], $B$=0.327, $SE$=0.130, $\beta$=0.502, $t(24)$=2.523, $p$=.019, two-tailed, mixed effects model with random intercept for participant (*30*). These findings suggest that infants inferred the rewards that the central agent placed over the targets from the cost the agent was willing to expend to reach these targets, and they therefore expected the agent to choose that target at test. Nevertheless, Experiment 1 does not show whether infants used the physical effort undertaken by the agent, or variables that merely correlate with effort (e.g. distance, speed), in their predictions.

[Fig 3 here]

To control for distance and speed of travel, Experiment 2 ($N$=24) used ramps of three different inclines to convey cost (Fig. 2B; Movie S2). On each familiarization trial, a target appeared on the top of one ramp, and the agent looked up the ramp and either climbed to the target or returned to its starting position. The agent climbed the shallow ramp and declined to climb the medium ramp for one target, and climbed the medium ramp and declined the steep ramp for the other target. The methods were otherwise the same as in Experiment 1. Consistent with our pre-specified directional prediction, infants again looked longer at the test events in which the agent approached the lower value target ($M$=30.94s, $SD$=13.31) than the higher value target ($M$=27.05s, $SD$=17.55) (Fig. 3), 95% CI [0.028, 0.472], $B$=0.250, $SE$=0.109, $\beta$=0.408, $t(24)$=2.294, $p$=.015, one-tailed (*30*). This finding further suggests that infants understand agents' actions in accord with abstract, general and interconnected concepts of cost and reward, but narrower explanations remain. In Experiments 1 and 2, the agent was confronted with an obstacle to its forward motion (a barrier or ramp), and the size of the obstacle covaried with the cost of the agent's action, requiring the agent to move further upward to attain the higher value target. Because infants become sensitive to the effects of gravity on objects' on inclined planes well before 10 months of age (*33*), they may learn that agents will move to greater heights or overcome higher obstacles for more rewarding targets, without invoking a more abstract representation of physical effort. Experiment 3 was undertaken to explore these interpretations.

In Experiment 3 ($N$=32), the agent was separated from each of the two targets during familiarization not by an obstacle but by a horizontal gap in the supporting surface (Fig. 2C, Movie S3). Infants first saw a ball roll off the edge of a narrow, medium, and wide gap, and shatter (Movie S6). During familiarization, these three trenches, requiring jumps of variable lengths and speeds but of equal durations and heights, were interposed between the agent and target; the agent moved to the edge of a trench, looked at the far side, and then jumped over a narrow trench for one target (and refused the medium trench), and a medium trench for the other target (and refused the widest trench). The methods were otherwise unchanged (Fig. 2E, Movie S5). The methods and analyses for Experiment 3 were preregistered at https://osf.io/k7yjt/ (*29*) and tested the same directional prediction as Experiment 2. Infants again looked longer at the lower value choice ($M$=23.05s, $SD$=13.58) relative to the higher value choice ($M$=17.47s, $SD$=10.69) (Fig. 3), 95% CI [0.020, 0.501], $B$=0.260, $SE$=0.119, $\beta$=0.403, $t(32)$=2.185, $p$=.018, one-tailed (*30*).

Regardless of whether an agent cleared higher barriers (Exp. 1), climbed steeper ramps (Exp. 2) or jumped wider gaps (Exp. 3) for one target over the other, infants expected the agent to choose that target at test. Across all experiments, infants looked longer at the lower value action ($M$=26.99s, $SD$=14.13) than the higher value action ($M$=21.64s, $SD$=13.94), 95% CI [0.139, 0.415], $B$=0.277, $SE$=0.070, $\beta$=0.424, $t(80)$=3.975, $p$<.001, one-tailed, mixed effects model with random intercepts for participant and experiment, supporting our general hypothesis that infants infer the values of agents' goals from the costs of their actions. Although past research had shown that infants represent the goal of an agent's action from observations of an agent's choices between two objects (*6*) and expect agents to give different emotional responses when agents complete versus fail to complete their goals (*31*), the present experiments provide evidence that infants develop ordinal representations of reward even when the number of choices and expressed emotions are equated across the actions and only the costs of the actions vary. Moreover, they show that infants do not simply attribute higher reward to goals that agents pursue for a longer duration or attain with greater frequency, because these variables were equated as well. The findings provide evidence for longstanding suggestions that infants represent physical cost as a continuous variable that agents seek to minimize (*3*, *13*): Infants make appropriate cost assessments even when the specific physical features that distinguished lower- from higher-cost actions, including the relative length, curvature, duration or speed of a motion trajectories, systematically varied. Together, Experiments 1-3 suggest that infants represent cost and reward as interconnected, abstract variables that they apply to a wide range of events.

The discovery that infants infer the rewards of goals from the costs of achieving them provides empirical support for the thesis that an abstract and productive system guides infants' analysis of agents and their actions (*3*, *17*, *19*). Specifically, we suggest that the cognitive machinery supporting infants' intuitive psychology includes a mental model both of how agents plan actions in the forward direction, in accord with maximizing their utilities (Eq. 1) (*23*), and a procedure for inverting this model, in accord with the computational framework of inverse planning (Eq. 2) (*4*). Applying this general framework to our specific experiments, we posit that infants have developed a model of action planning prior to the experiment: they assume that agents value some goal objects more than others, and to engage in costlier actions to achieve goals with higher reward. When the infants see the agent take costlier actions to arrive at one target than at another, they invert this model to infer the relative reward of the two targets to that agent. Then when they see the agent flanked by the two targets in a situation where costs are equal, they apply their knowledge of the targets' relative value to the agent to run their planning model for that agent forward, predicting the target that it will approach. We have implemented this hypothesis in a computational model that accounts not only for the findings of the present experiments but also for a range of past studies of early action understanding (*6*, *7*, *13*). Furthermore, we compared this model to an array of simpler models that focus only on relative costs or rewards in isolation, or on particular cues to effort or value. We find that the only the full model with abstract variables for costs and rewards can account for all of the findings (Fig. S3; see Supplementary Material for details).

The present studies raise key questions for future research. First, the cognitive architecture underlying infants' assessment of cost remains to be explored. Our experiments suggest that infants are responding to an abstract notion of cost, rather than specific physical path features such as vertical motion (controlled for in Exp. 3), horizontal motion (controlled for in Exp. 1), or raw path length (controlled for in Exp. 2). We do not know, however, whether infants represent the abstract costs of actions by drawing on a concept of experienced effort or exertion within the domain of naïve psychology, or by leveraging an intuitive concept of force or work done (i.e. the integral of force applied over a path) from the domain of naïve physics (*34*, *35*), or perhaps both. Next, our experiments investigated only one class of goal states and target-directed actions, leaving open the breadth and generality of infants' intuitive psychology. In particular, cost can be defined in terms of work or effort to produce physical forces, but there are other kinds of costs: Agents could consider variables like the mental effort of planning (*36*, *37*) and the risks of choosing certain actions, neither of which involves applications of force. It is an open question whether these other variables trade off against reward in infants' intuitive psychology the way that physical work or effort does. Lastly, our studies do not speak to the origins of these abilities. Although 10-month-old infants cannot perform the actions from our experiments or communicate with others about them, their productive system for reasoning about costs and rewards may arise through their experiences observing the actions of other agents or performing actions within their repertoire, such as lifting their arms or balancing their bodies against the force of gravity. Alternatively, this system of intuitive psychology may guide infants' action understanding from the beginning. Testing these possibilities would address fundamental questions concerning the nature, origins, and interrelations between our intuitive psychology and intuitive physics.

However these questions are answered, the present study suggests that our propensity to understand the minds and actions of others in terms of abstract, general and interrelated concepts begins early. Before human infants learn to walk, leap and climb, they leverage mental models of agents and actions—forward models of how agents plan, and inverse models for working backwards from agents' actions to the causes inside their minds.

**References and Notes:**


1. F. Heider, M. Simmel, An experimental study of social behavior. *Am. J. Psychol.* **57**, 243–259 (1944).
2. D. C. Dennett, *The Intentional Stance* (The MIT Press, London, 1987).
3. G. Gergely, G. Csibra, Teleological reasoning in infancy: The naïve theory of rational action. *Trends Cogn. Sci.* **7**, 287–292 (2003).
4. C. L. Baker, R. Saxe, J. B. Tenenbaum, Action understanding as inverse planning. *Cognition*. **113**, 329–349 (2009).
5. R. Saxe, T. Tzelnic, S. Carey, Five-month-old infants know humans are solid, like inanimate objects. *Cognition*. **101**, B1–B8 (2006).
6. A. L. Woodward, Infants selectively encode the goal object of an actor's reach. *Cognition*. **69**, 1–34 (1998).
7. G. Gergely, Z. Nádasdy, G. Csibra, S. Bíró, Taking the intentional stance at 12 months of age. *Cognition*. **56**, 165–193 (1995).
8. K. H. Onishi, R. Baillargeon, Do 15-month-old infants understand false beliefs? *Science*. **308**, 255–8 (2005).
9. Y. Luo, R. Baillargeon, Do 12.5-month-old infants consider what objects others can see when interpreting their actions? *Cognition*. **105**, 489–512 (2007).
10. Y. Luo, S. C. Johnson, Recognizing the role of perception in action at 6 months. *Dev. Sci.* **12**, 142–149 (2009).
11. P. Muentener, S. Carey, Infants' causal representations of state change events. *Cogn. Psychol.* **61** (2010), pp. 63–86.
12. A. E. Skerry, S. E. Carey, E. S. Spelke, First-person action experience reveals sensitivity to action efficiency in prereaching infants. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 18728–33 (2013).
13. S. Liu, E. S. Spelke, Six-month-old infants expect agents to minimize the cost of their actions. *Cognition*. **160**, 35–42 (2017).
14. A. L. Woodward, Infants' grasp of others' intentions. *Curr. Dir. Psychol. Sci.* **18**, 53–57 (2009).
15. M. Paulus, Action mirroring and action understanding: an ideomotor and attentional account. *Psychol. Res.* **76**, 760–7 (2012).
16. C. M. Heyes, C. D. Frith, The cultural evolution of mind reading. *Science*. **344**, 1243091 (2014).
17. E. S. Spelke, K. D. Kinzler, Core knowledge. *Dev. Sci.* **10**, 89–96 (2007).
18. J. B. Tenenbaum, C. Kemp, T. L. Griffiths, N. D. Goodman, How to grow a mind: statistics, structure, and abstraction. *Science*. **331**, 1279–1285 (2011).
19. R. Baillargeon, R. M. Scott, L. Bian, Psychological Reasoning in Infancy. *Annu. Rev. Psychol.* **67**, 159–186 (2016).
20. D. Bernoullli, Exposition of a new theory on the measurement of risk. *Econometrica*. **22**, 23–36 (1954).

21. J. S. Mill, *Utilitarianism* (Oxford University, 1863).
22. J. Bentham, *An Introduction to the Principles of Morals and Legislation* (Clarendon Press, 1879).
23. J. Jara-Ettinger, H. Gweon, L. E. Schulz, J. B. Tenenbaum, The Naïve Utility Calculus: Computational Principles Underlying Commonsense Psychology. *Trends Cogn. Sci.* **20**, 589–604 (2016).
24. C. G. Lucas *et al.*, The child as econometrician: A rational model of preference understanding in children. *PLoS One*. **9** (2014), doi:10.1371/journal.pone.0092160.
25. A. Jern, C. Kemp, A decision network account of reasoning about other people's choices. *Cognition*. **142**, 12–38 (2015).
26. J. Jara-Ettinger, H. Gweon, J. B. Tenenbaum, L. E. Schulz, Children's understanding of the costs and rewards underlying rational action. *Cognition*. **140**, 14–23 (2015).
27. Y. Luo, R. Baillargeon, Can a self-propelled box have a goal? - Psychological reasoning in 5-month-old infants. *Psychol. Sci.* **16**, 601–608 (2005).
28. G. Csibra, G. Gergely, S. Bíró, O. Koós, M. Brockbank, Goal attribution without agency cues: The perception of "pure reason" in infancy. *Cognition*. **72**, 237–267 (1999).
29. Detailed methods and analyses from the experiments reported in this paper are available in the Supplemental Materials.
30. Prior to conducting Exp. 1, we were unsure whether infants would express their expectations by longer or shorter looking at the unexpected event, and so we pre-specified a two-tailed alpha. After Exp. 1, however, we had strong reason to believe that infants would look longer at the test actions that they found less probable, and so we pre-specified (for Exp. 2) and pre-registered (for Exp. 3) one-tailed tests and report these numbers in the text. However, results of all three experiments were also statistically significant by two-tailed tests, with $p$=.031 and $p$=.036 for Exp. 2 and 3, respectively.
31. A. E. Skerry, E. S. Spelke, Preverbal infants identify emotional reactions that are incongruent with goal outcomes. *Cognition*. **130**, 204–216 (2014).
32. G. Csibra, M. Hernik, O. Mascaro, D. Tatone, M. Lengyel, Statistical treatment of looking-time data. *Dev. Psychol.* **52**, 521–536 (2016).
33. I. K. Kim, E. S. Spelke, Infants' sensitivity to effects of gravity on visible object motion. *J. Exp. Psychol. Hum. Percept. Perform.* **18**, 385–393 (1992).
34. E. Téglás *et al.*, Pure reasoning in 12-month-old infants as probabilistic inference. *Science (80-. ).* **332**, 1054–9 (2011).
35. T. D. Ullman, E. Spelke, P. Battaglia, J. B. Tenenbaum, Mind games: Game engines as an architecture for intuitive physics. *Trends Cogn. Sci.* **xx**, 1–17 (2017).
36. A. Shenhav *et al.*, Toward a Rational and Mechanistic Account of Mental Effort. *Annu. Rev. Neurosci.* **40**, 99–124 (2017).
37. W. Kool, S. J. Gershman, F. A. Cushman, Cost-Benefit Arbitration Between Multiple Reinforcement-Learning Systems. *Psychol. Sci.* **28**, 1321–1333 (2017).
38. B. Foundation, Blender (2016), (available at blender.org/download).
39. J. Pinto, XHAB64 (1995).
40. R. M. Casstevens, jHab: Java Habituation Software (2007).
41. R. D. C. Team, R: A language and environment for statistical computing (2015), (available at https://www.r-project.org/).
42. D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67** (2015), doi:10.18637/jss.v067.i01.

43. R. Nieuwenhuis, M. te Grotenhuis, B. Pelzer, Influence.ME: Tools for detecting influential data in mixed effects models. *R J.* **4**, 38–47 (2012).
44. H. Wickham, *ggplot2: Elegant graphics for data analysis* (Springer-Verlag, 2009).
45. G. Gergely, H. Bekkering, I. Király, Rational imitation in preverbal infants. *Nature*. **415**, 755 (2002).
46. A. N. Meltzoff, Understanding the intentions of others : Re- enactment of intended acts by 18-month-old children. *Dev. Psychol.* **31**, 838–850 (1995).
47. Y. Mou, J. M. Province, Y. Luo, Can infants make transitive inferences? *Cogn. Psychol.* **68**, 98–112 (2014).
48. T. Kushnir, F. Xu, H. M. Wellman, Young children use statistical sampling to infer the preferences of other people. *Psychol. Sci. a J. Am. Psychol. Soc. / APS*. **21**, 1134–1140 (2010).
49. G. Csibra, S. Bíró, O. Koós, G. Gergely, One-year-old infants use teleological representations of actions productively. *Cogn. Sci.* **27**, 111–133 (2003).
50. A. Jern, C. Kemp, in *Proceedings of the 36th {A}nnual {C}onference of the {C}ognitive {S}cience {S}ociety*, P. Bello, M. Guarini, M. McShane, B. Scassellati, Eds. (2014).
51. J. . Hamlin, T. Ullman, J. Tenenbaum, N. Goodman, C. Baker, The mentalistic basis of core social cognition: experiments in preverbal infants and a computational model. *Dev. Sci.* **16**, 209–226 (2013).
52. S. Russell, P. Norvig, *Artificial Intelligence: A Modern Approach, Third edition* (2014).
53. R. Sutton, A. Barto, Reinforcement Learning: An Introduction. *MIT Press. Cambridge, Massachusetts* (1998).
54. N. D. Goodman, V. K. Mansinghka, D. M. Roy, K. Bonawitz, J. B. Tenenbaum, Church: a language for generative models. *Uncertain. Artif. Intell.* (2008).

**Acknowledgements**: All authors wrote the paper, S.L. designed the experiments and collected and analyzed the data under the direction of E.S., and T.U. built the computational models under the direction of J.T. This material is based upon work supported by the Center for Brains, Minds and Machines (CBMM), funded by National Science Foundation STC award CCF-1231216, and by a National Science Foundation Graduate Research Fellowship under grant DGE-1144152. We thank the families who volunteered to participate, and A. Aguirre, R. Guzman, C. Kerwin, N. Kalra, and H. Tarr for research assistance. All data, analyses, and materials from the experiments and models reported in this paper are archived on the Open Science Framework at https://osf.io/mykph/.

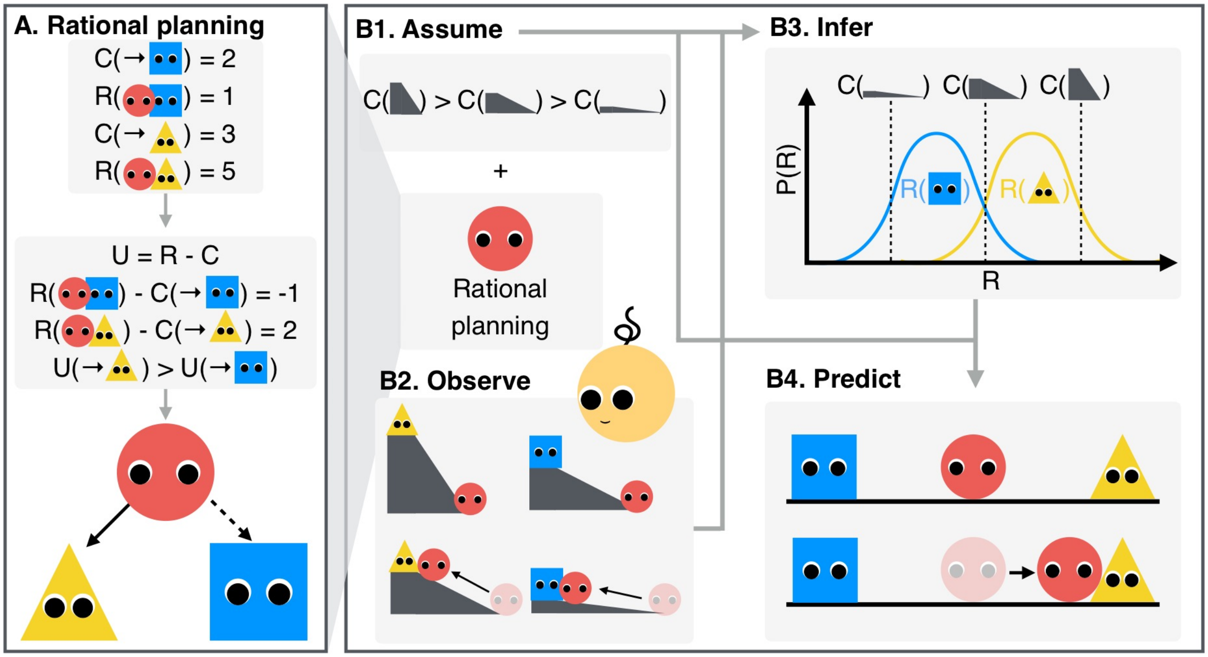


**Fig. 1.** A schematic of our computational model. The forward direction (A) defines the agent as a rational planner that calculates the utilities of different actions from their respective costs and rewards, and then selects an action stochastically in proportion to its utility. In this case, the overall utility for approaching Triangle is higher than for approaching Square, so the central agent (Circle) will likely choose Triangle over Square. An observer (B1) 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.

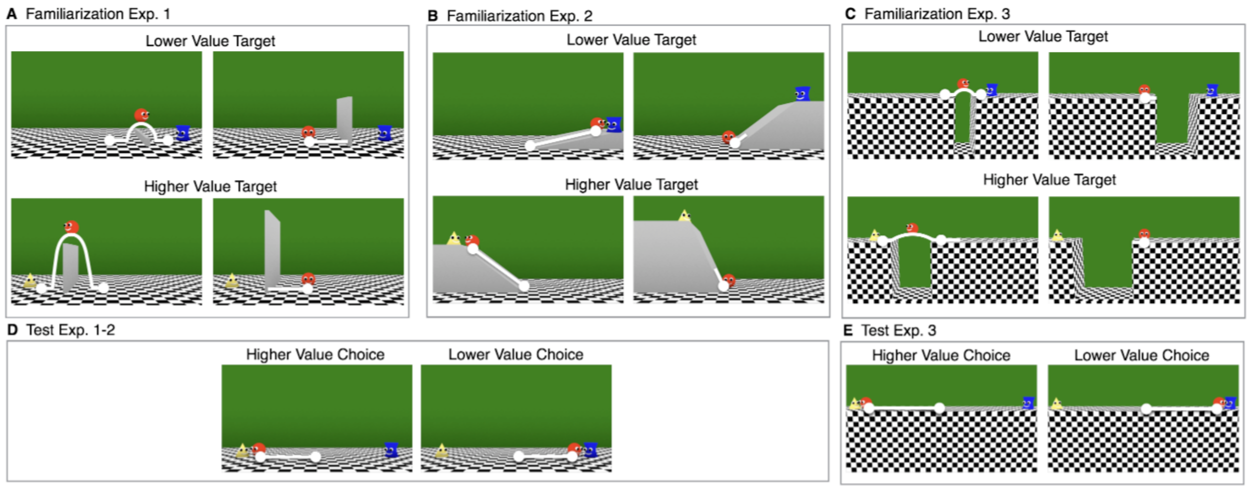


**Fig. 2.** Structure of Experiments 1-3. During familiarization (A-C), the central agent (Circle) accepted a low and refused a medium cost for the lower value target (in this case, Square), and accepted a medium and refused a high cost for the higher value target (Triangle). Other than the sizes of the barriers, ramps, and trenches, and the consequent trajectories of motion, the pairs of events displaying approach or refusal of approach to the two targets were identical. At test (D-E), the agent 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.

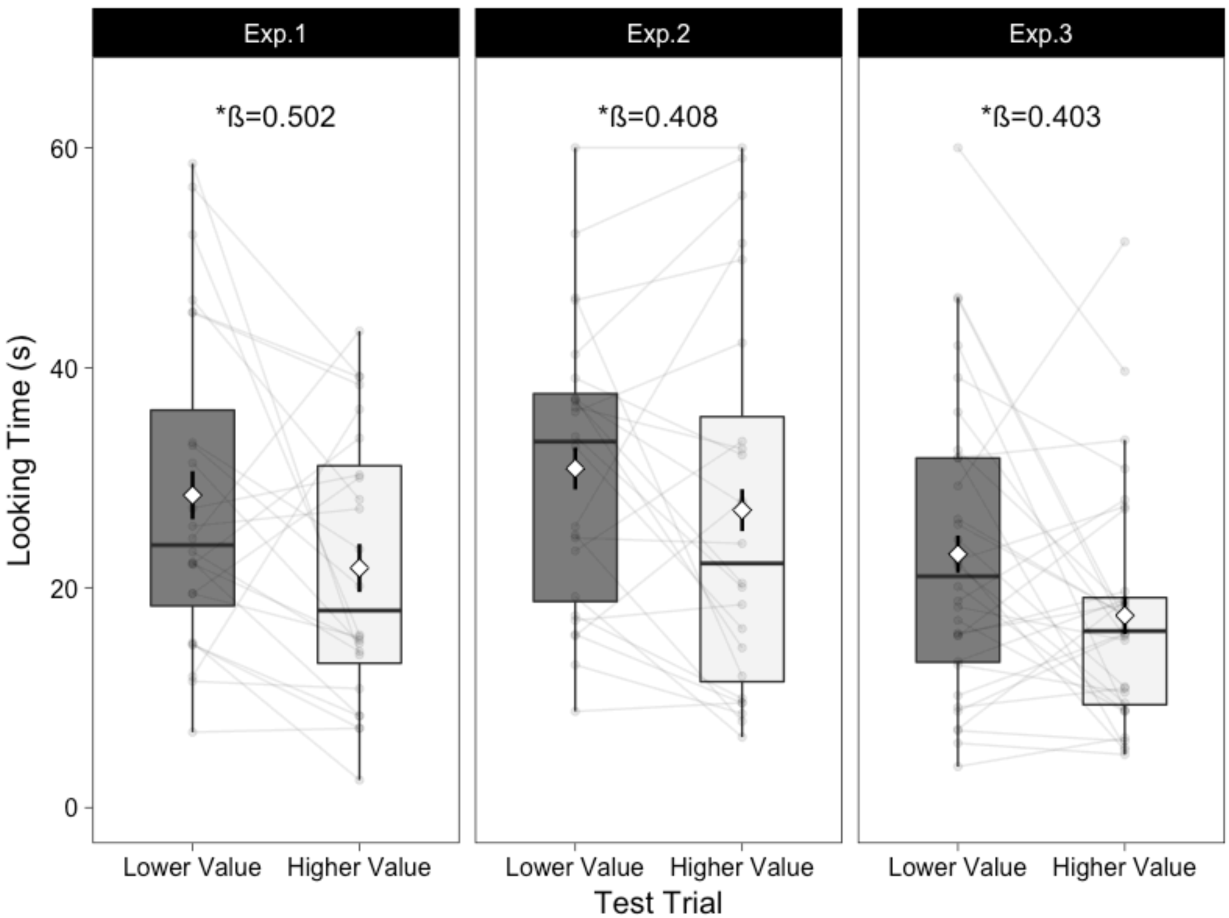


**Fig. 3.** Boxplots of average looking time towards the higher and lower value choice during test in Experiments 1-3. White diamonds indicate means, with error bars indicating within-subjects standard errors. Horizontal lines indicate medians, boxes indicate middle quartiles, and whiskers indicate points within 1.5 times the interquartile range from the upper and lower edges of the middle quartiles. Light grey points connected across boxes indicate looking times from individual participants. Beta coefficients indicate effect sizes in standard deviations, and asterisks indicate significance relative to pre-specified (Experiments 1-2) and pre-registered (Experiment 3) alphas (*<.05). See text and SM for statistical analyses.