The Naïve Utility Calculus: Computational Principles Underlying Commonsense Psychology

Citation: Jara-Ettinger, Julian et al. "The Naïve Utility Calculus: Computational Principles Underlying Commonsense Psychology." Trends in Cognitive Sciences 20, 8 (August 2016): 589–604 © 2016 Elsevier Ltd

As Published: http://dx.doi.org/10.1016/j.tics.2016.05.011

Publisher: Elsevier

Persistent URL: http://hdl.handle.net/1721.1/113204

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

The naïve utility calculus: Computational principles underlying commonsense psychology

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

[a] Department of Brain & Cognitive Sciences, MIT, Cambridge, MA 02139

[b] Department of Psychology, Stanford University, Stanford, CA 94305

Abstract

We propose that human social cognition is structured around a basic understanding of ourselves and others as intuitive utility maximizers: From a young age, humans implicitly assume that agents choose goals and actions to maximize the rewards they expect to obtain relative to the costs they expect to incur. This "naïve utility calculus" lets both children and adults observe others' behavior and infer their beliefs and desires, their longer-term knowledge and preferences, and even their character: who is knowledgeable or competent, who is praiseworthy or blameworthy, who is friendly, indifferent or an enemy. We review studies providing support for the naïve utility calculus, and we show how it captures much of the rich social reasoning humans engage in from infancy.

Commonsense Psychology

Theories of decision-making have been at the heart of psychology since the field's inception, but only recently has the field turned to the study of how humans – especially the youngest humans – think humans make decisions. When we watch someone make a choice, we explain it in terms of their goals, preferences, personalities, and moral beliefs. This capacity – our commonsense psychology – is the cognitive foundation of human society. It lets us share what we have and know, with those from whom we expect the same in return, and it guides how we evaluate those who deviate from our expectations.

The representations and inferential power underlying commonsense psychology trace back to early childhood – before children begin kindergarten, and often even in infancy. Work on how children reason about other agents' goals [1-8], desires [9-11], beliefs [12-18], and pro-social behavior [19-29] has advanced our understanding of what in our commonsense psychology is at work in early infancy [30-32] and what develops [16-17,33-35]. Nonetheless, major theoretical questions remain unresolved. What computations underlie our commonsense psychology, and to what extent are they specific to the social domain? Are there a small number of general principles by which humans reason about and evaluate other agents, or do we instead learn a large number of special case rules and heuristics? To what extent is there continuity between the computations supporting commonsense psychology in infancy and later ages? Is children's social-cognitive development a progressive refinement of a computational system in place from birth, or are there fundamentally new computational principles coming into play?

In this article we advance a hypothesis that offers answers to each of these questions, and provides a unifying framework in which to understand the diverse social-cognitive capacities we see even in young children. We propose that human beings, from early infancy, interpret others' intentional actions through the lens of a naïve utility calculus: that is, people assume that others choose actions to maximize utilities -the rewards they expect to obtain relative to the costs they expect to incur. The naïve utility calculus can be made precise computationally and tested quantitatively (Box 1). Embedded in a Bayesian framework for reasoning under uncertainty, and supplemented with other knowledge children have about the physical and psychological world (e.g. knowledge about objects, forces, action, perception, goals, desires, and beliefs), the naïve utility calculus supports a surprisingly wide range of core social-cognitive inferences and it persists stably in some form through adulthood, guiding the development of social reasoning even as children's thinking about others undergoes significant growth.

Figure 1 illustrates some of the basic social intuitions that go beyond goal attribution which the naïve utility calculus aims to explain. These examples illustrate the role costs and rewards play in commonsense psychology, but they are not specific to agents choosing fruits on shelves; they apply to a wide range of situations in which intentional agents of any sort (child, adult, animated ball) interact with each other and move toward, reach for, or manipulate objects. We focus our discussion on behaviors where even young children can immediately grasp the costs and rewards involved. The naïve utility calculus likely applies to more abstract situations as well, but its application may be complex in ways we do not consider here (e.g., cases where cultural norms are in play). Although we focus on intentional behavior (as opposed to habits, reflexes, accidents, etc.) some of the most revealing choices are decisions not to act; our proposal aims to account for these as well.

The ideas behind this naïve utility calculus have a long history, tracing back to classical philosophers like Adam Smith [36] and John Stuart Mill [37]. Its formulation as an intuitive theory was anticipated in some form by pioneers in social cognition Fritz Heider [38], Harold Kelley [39], and Roger Brown [40], but with the development of new computational cognitive modeling tools these ideas can be formulated and tested more precisely [41-46].

Critically, the naïve utility calculus is not a scientific account of how people act; it is a scientific account of people's intuitive theory of how people act. These two notions may diverge – indeed, the mathematics of utility theory was originally proposed by early economists [47] but fails to predict actual human behavior in many important economic contexts ([48, 49]). But this does not mean the naïve utility calculus is not in some sense a reasonable and useful model of human behavior. In physics, our intuitive theory is oversimplified with respect to how the physical world actually works [50], yet it still helps us navigate everyday life because it tends to support accurate predictions on the spatial and temporal scales that matter most to us [51]. Similarly, the naïve utility calculus does not require that agents actually compute and maximize fine-grained expected utilities in order to be a useful guide in many everyday social situations.

In the remainder of this article we first describe the crucial ideas of the naïve utility calculus in their simplest, most ideal form. Next we move on to more nuanced features of the intuitive theory necessary to apply it to real-world decision making. We follow by reviewing studies that directly test the proposal, as well as the broader literature on goal-directed action, sampling-sensitive and preference judgments, communication, pedagogy, and social and moral evaluation that can be explained by our framework. We conclude with a discussion of how the naïve utility calculus relates to accounts of first-person decision-making, and the proper relationship between intuitive and scientific theories of intentional action.

Naïve utility theory: Agents as utility-maximizers

Formally, we propose that the naïve utility calculus consist of a theory or a generative model, which, embedded in a Bayesian framework, supports predictions about future behaviors (setting the costs and rewards and deriving the resulting actions) and inferences about the causes of observed behaviors (finding, through Bayes' rule, the costs and rewards that can generate the observed actions). A formal description of the proposal is presented in Box 1.

The generative model specifies how costs and rewards determine agents' behavior. When agents decide how to act (e.g., whether to pursue a goal or which goal to pursue), they estimate the expected utility of each goal. Each goal's utility is calculated by estimating the rewards the agent would obtain if she completed the goal, and subtracting the cost she would need to incur to complete it. Through this process, agents build a utility function that maps possible plans onto expected utilities. Agents then pursue the plan with the highest positive utility. As such, agents are only willing to pursue plans where the rewards outweigh the costs, and if a plan has negative utility, the agent will be unwilling to act upon it, even in the absence of alternatives.

By assuming that agents behave in accordance with the generative model, observers can work backwards to infer the set of costs and rewards most likely to have generated the observed behavior. In line with previous work (e.g., [41]), we propose that these reverse inferences can be understood as a kind of Bayesian inference (Box 1).

The naïve utility calculus (an intuitive theory of agents as a generative model and a way to reverse this model through Bayesian inference) makes some key predictions about how humans reason about others' behavior, some of which are shown in Figure 2. These predictions not only involve people's qualitative judgments but also their confidence, supporting inferences about the ambiguity of exact rewards and costs underlying others' behaviors. For simplicity, here we only consider the rewards associated with outcomes and the costs associated with sequences of actions; as we note below, however, an outcome can be costly and the sequence of actions be rewarding, too.

Real world reasoning with a naïve utility calculus

Reasoning about decision-making in the real world, however, has several complications that the idealized naïve utility calculus cannot handle. These complications reveal more sophisticated aspects or the naïve utility calculus that give it traction and point to ways in which commonsense psychology may develop (Box 2).

First, agents do not always know their costs and rewards when deciding how to act. As such, agents do not maximize true utilities, but expected utilities. In familiar scenarios, agents should make accurate estimates. Most people, for instance, can estimate their costs for walking a block and their rewards from eating a cookie. However, agents often pursue novel outcomes in novel ways. In these contexts, it is critical that observers understand that agents act based on the expected rather than true costs and rewards (Figure 3). Observers should be less likely to infer that agents' choices are stable if the agent might have been ignorant or mistaken about the true costs or rewards of her actions, as will often happen for agents who are inexperienced with the rewards they are choosing (Figure 1g and 1h).

Moreover, agents have to estimate their own expected utilities, and these estimates may be inexact (see Box 1). Intuitively, when two plans have very different expected utilities, it is easy to identify the better plan. However, when plans have similar expected utilities, agents may find it more difficult to decide which is best – even apart from any uncertainty in their basic costs and rewards. This assumption provides flexibility in observers' inferences, softening the assumption that choices unambiguously reveal the highest expected utility. It also allows observers to infer agents' costs and rewards from the dynamics of their decision-making: agents are more likely to deterministically choose one plan over another when their utilities are very different, and more likely to oscillate between their choices when the utilities are similar.

Second, costs and rewards are not objective properties of the physical world, but subjective experiences that vary across agents. Some people find walking more difficult than others, and some people like cookies more than others. However, the structure of costs and rewards also has an agent-invariant structure. Two cookies are better than one and longer distances are costlier to travel than short ones. These individual differences may be observable or may have to be inferred as part of explaining an agent's actions, similar to classic attribution theories [52]. By integrating both agent-invariant (objective) and agent-dependent (subjective) aspects of costs and rewards, the naïve utility calculus allows learners to parcel out known agent-invariant contributions to how an agent acts in a given situation and thereby infer latent costs and rewards that differ across agents.

Third, the content of costs and rewards goes far beyond physical actions and outcomes. In social situations, an agent's costs and rewards can depend recursively on their expectations about another agent's costs and rewards. If someone is motivated to help, her rewards depend not only on her own utilities, but also on promoting the other person's utilities, or diminishing them if she is motivated to hinder [53]. Likewise, acting against what you know another agent wants you to do may impose a cost. By integrating an agent's own first-order (self-interested) costs and rewards with that agent's second-order appreciation of others' costs and rewards, the naïve utility calculus allows observers to make inferences about the nature and extent of others' prosocial or altruistic tendencies.

Finally, behaviors can have more than one cost-reward decomposition. When agents act they may incur costs for the actions and obtain rewards for the outcome; they may obtain rewards for the actions and incur costs for the outcome; or they may obtain rewards for both the actions and the outcome. The naïve utility calculus in its most general form supports all of these representations. However, this flexibility implies that behaviors have multistable cost-reward decompositions. As in other domain of cognitions (e.g., [54]), and consistent with the Bayesian framework [55], this challenge can be solved through an appropriate inductive bias or prior. We assume that as a default, people most naturally parse plans in terms of costly actions and rewarding outcomes, as shown in Figure 1. Other decompositions of costs and rewards can be invoked when these favored explanations are unable to account for the behavior (e.g., ascribing rewards to actions in themselves; [56]).

Evidence for the naïve utility calculus

Our empirical work provides several lines of evidence that the naïve utility calculus supports early social reasoning. In one series of experiments [57] we found that when five-year-olds learn an agent's costs and choices, they infer a reward function that guarantees that the agent maximized her utilities. We showed children a puppet who chooses crackers over cookies when both items are equidistant, but cookies over crackers when the cookies are closer (Figure 4a). If children equate choice with preference, they should think the puppet likes crackers and cookies equally; instead, our results show that children integrate the puppet's choices with cost information and recognize that the puppet prefers crackers (i.e., the item chosen when the costs were matched; Figure 4c). Similarly, when children observe agents' choices whose rewards are known, they infer a cost function that guarantees utility maximization. We showed children one puppet who liked crackers more than cookies and another puppet who liked them both equally. We then put the cookies on a low box and the crackers on a high box. Both puppets chose the cookies (Figure 4b). When asked which puppet couldn't climb, children chose the puppet with the strong preference even though neither puppet even attempted to climb (Figure 4c). Further experiments also showed that children understand how different agents can incur different costs (i.e., costs vary across agents) even when taking identical actions.

The naïve utility calculus implies that agents who are ignorant about the costs and rewards of actions should be more likely to make poor choices and change their minds (and conversely, that agents who make poor choices and change their minds are likely ignorant about costs and rewards). We introduced four-year-olds to two puppets, both of whom reached for and chose a rambutan over an African cucumber (see Fig 1. H, and Fig. 3). One puppet then said "yuck" (or in a separate experiment, changed her mind). Children were asked which puppet knew all about these fruits before and which had never seen them before. Children successfully identified the naïve agent (and conversely, if they knew which agent was knowledgeable and which naïve, they could guess who said "yuck"). Children were able to draw similar inferences with respect to inferences about agents' costs (see Fig 1. g; [58]).

In another set of experiments [59], we showed that the naïve utility calculus supports toddlers' social evaluations. We showed two-year-old children two puppets making a toy play music; one puppet was able to make the toy play music on the first try (low cost) while the other took several attempts (high cost). At baseline, toddlers preferred to play with the more competent agent and judged him to be nicer. When both puppets refused to help the parent activate the toy, toddlers continued to prefer the more competent agent but now judged that the less competent agent was nicer (See Figure 4d and 4e). Consistent with the naïve utility calculus, these results suggest that two-year-olds can infer an agent's motivation to help (her subjective rewards) given information about her costs and, like adults, are more likely to exonerate agents for whom helping is costly than those who are simply unmotivated to be helpful.

The naïve utility calculus as a unifying framework for social cognition

Beyond these studies that directly test the predictions in Figure 1, the naïve utility calculus has implications for a wide array of other phenomena in social cognitive development. As noted, researchers have looked extensively at children's intuitions about agents' goal-directed actions, desires and beliefs, pro-social behavior, and teaching and learning from others. Each of these aspects of social cognition has typically been treated as a separate problem, and explored through different paradigms. However, findings in many of these areas can be unified under the assumption that humans predict and explain behavior through a naïve utility calculus, as we illustrate below.

Goal-directed action

A large body of work in cognitive development suggests that even infants expect agents to complete their goals as efficiently as possible [2-3,60-65]. If for instance, infants are habituated to one agent hopping over a barrier to reach another agent, infants look longer when the agent continues to hop in the absence of a barrier than when she moves in a straight line [2,65].

Theses inferences have been explained by the hypothesis that infants adopt a "teleological stance" [3], a non-mentalistic representation of behavior where agents are assumed to move efficiently towards goal-states, subject to situational constraints. The teleological stance is thought to underlie infants' earliest forms of reasoning about agents and to serve as the basis for mentalistic representations that emerge later in life. The teleological stance is compatible with the naïve utility calculus: if agents maximize utilities, they should incur the minimum costs necessary to obtain rewards. However, the naïve utility calculus expands on the teleological stance by explaining how agents select their goals, and by explaining how objective (e.g., walls) and subjective (e.g., competence) constraints not only influence goal-completion, but also goal-formation. Related ideas have been explored [45,66], although not with the same focus on cost-reward tradeoffs in childhood as we emphasize here.

Is it possible that infants merely expect agents to take the shortest possible path to a goal, without an abstract representation of costs or an expectation that agents should minimize them? Several studies suggest that infants represent efficiency in terms of relative costs that go beyond simply computing the length of the path. Southgate et al. [64] showed that infants appear to expect actions with fewer number of steps to be performed, over actions that take more steps or more time. Gergely et al. [67] showed infants an actor who used their head to light up a toy when their hands were either free or occupied. Infants themselves were more likely to imitate the head action in the hands-free condition compared to the hands-occupied condition, suggesting they inferred the actor had a specific intention (indicating a source of strong reward) to use their heads only when that was clearly the more costly of available alternative actions. Together, these findings suggest that infants' expectation for efficient action may be driven by an abstract notion of cost-minimization. Nevertheless, experiments that directly pit a path's simplicity, straightness, length, time and energy costs against each other are needed to reveal if a general metric of cost minimization is at work an infancy, or if it arises later, building on top of some more limited, primitive notion of action efficiency.

More generally, a number of studies suggest that infants believe that the ability to perform effortful, high cost actions in the service of salient or plausible goals is the special provenance of agents (and only agents). Abilities attributed to agents (but not to objects or physical forces) include the ability to engage in self-generated movement [4,68-69], the ability to resist gravity [70], the ability to cause objects to move or change state [71-72], the ability to create order [73], the ability to generate patterns [74], and the ability to spontaneously and non-deterministically cause changes in the world [75-76].

Such studies provide evidence that infants have intuitions about the costs of agents' actions. Other work suggests that infants also understand the rewards of goal-directed actions. Ten-month-olds appear surprised when an agent expresses a negative emotion following a completed (versus failed) goal [77], suggesting that they expect agents to find goal-completion rewarding. Ten-month-olds also attribute a preference to an agent who consistently chooses one goal over another [11] suggesting that infants understand that agents can find some goals more rewarding than others. By 18-months, children also understand that different agents can find the same goal more or less rewarding [10].

Collectively these results suggest that at least many key prerequisites to a naïve utility calculus emerge early in development: an expectation that agents act efficiently in the sense of acting to maximize rewards relative to costs, an expectation that agents (and only agents) can perform effortful actions in the service of goals, and an expectation that agents experience subjective rewards consistent with goal outcomes.

Sampling and preferences

Infants as young as six months expect randomly sampled sets, but not deliberately selected sets, to be representative of the population from which they are drawn [78-82]. This sensitivity to the sampling process supports learning properties of novel objects [83] and the scope of the meaning of novel words [84-85]. For instance, Gweon et al [83] showed 15-month-old infants a box full of blue and yellow toys and an agent taking out three blue toys to demonstrate that they all share some hidden property (e.g., they squeak). Infants appeared to expect all toys to share the hidden property when the blue balls were common (suggesting that the agent sampled three blue balls by chance), but not when they were rare (suggesting that the agent sampled three blue balls selectively). In the absence of a clear purpose behind an agent's sampling actions, infants attribute preferences [9,11]. For instance, if an agent pulls three frogs in a row from a box that contains mostly ducks, 20-month-olds infer that the agent prefers frogs to ducks; they do not infer this if the box contains more frogs than ducks or if the box contains only frogs.

The intuitions underlying toddlers' and infants' sensitivity to the sampling process can be explained through the naïve utility calculus. This is easy to see if we imagine unpacking a population of objects into a generic spatial configuration where objects are randomly distributed in space, with some closer and others further from an agent, and hence less costly or more costly for the agent to reach (see Figure 5). If all the objects in a box are equally rewarding, then agents should minimize costs by taking the objects that are the easiest to reach, generating a sample representative of the population. However, if one type of object is more rewarding than the others, then the agent should selectively draw that kind of object even if it is more costly to obtain, generating a biased sample. Reversing these inferences, if an agent generates a sample that could have been obtained simply by minimizing costs, her actions provide no reason to think that some objects are more rewarding than others. However, if generating the sample required the agent to perform costly actions (in time, effort, and attention), the rare objects must have been more rewarding.

Communication and pedagogy

The naïve utility calculus also provides a principled explanation for how the assumptions underlying pedagogical communication emerge. If a teacher shares information, the reward from sharing must exceed the cost for teaching. As such, in small and simple domains (e.g., a toy with just a few functions) where the cost of sharing information is negligible, agents should share all the information necessary for the learner to draw accurate inferences, and learners can use this expectation to make inferences accordingly. Consistent with this expectation, children assume that teachers share all relevant information in simple domains [86], and when a teacher demonstrates only one of many functions of a toy, children rate the teacher poorly and mistrust his subsequent teaching [87]. These inferences should be cost-sensitive, however: learners' expectation that informants will communicate all relevant information should be weaker when the costs are higher. Consistent with this, children prefer exhaustive informants when costs are low but prefer informants who provide only information sufficient for a good inductive inference when costs are high [88]. In more complex domains, more complex inferences are warranted. The naïve utility calculus makes the untested predictions that observers should be less surprised if teachers fail to provide exhaustive evidence about a toy with many functions than a toy with only a few. Similarly, if a toy has many equally rewarding functions but some are costlier to demonstrate than others, observers should be less surprised if the teacher fails to share high-cost information than low-cost information.

As noted earlier, the costs and rewards of pedagogy crucially can have recursive components: In addition to the teacher's intrinsic reward for teaching, her utilities for sharing some information may depend on how rewarding it is for the learner to learn it, and how costly the learner's different learning options are. Consistent with this, a number of studies suggest that very young children go out of their way to communicate information that is currently unknown to the learner [89-90], relevant to the learner's goal [91], or difficult for the learner to discover by herself [92-93].

Finally, in linguistic communication more broadly, the classic Gricean maxims – that speakers communicate things that are relevant to the conversation (maxim of relation), and they provide all the information needed (maxim of quantity) in a manner that is truthful (maxim of quality) and clear (maxim of manner). [94] - are central in pragmatic inferences for both adults and children [95-96] and can be derived from the naïve utility calculus. Minimizing utterance length (communication costs) while maximizing information transfer to the listener (communicative rewards) can be seen as optimizing an overall utility function trading off these costs and rewards.

Social and moral reasoning

Many studies have suggested that social evaluation emerges in the first year of life, with infants preferring agents who help others achieve their goals to those who hinder those goals [19,23]. Moreover, infants' evaluations are transitive (they prefer agents who hinder hinderers and help helpers; [97-98]) and they only positively evaluate agents if they helped intentionally [21] and did so with knowledge of the recipients' preferences [22]. Such studies are consistent with a naïve utility calculus: in every case, the helper or hinderer takes costly actions (i.e., goes out of his/her way to intervene), supporting the inference that the goal (helping or hindering the other) must be rewarding. Moreover, these studies suggest that infants may already understand that agents' utilities can go beyond including their individual costs and rewards and also integrate others' costs and rewards. As noted, our own work suggests that toddlers also use agents' relative costs to distinguish their motivations: if someone refuses to help when helping is costly, two-year-olds think she is nicer than a more competent agent who refuses to help at low cost [59].

The naïve utility calculus has many other, untested, implications for social evaluation. Consider for instance, that agents who underestimate rewards or costs may be more liable to abandon plans or commitments, with consequences for how others judge them and whether they trust them in the future. It is also noteworthy that there is a special category of moral blame ("exploitation") for those who knowingly take advantage of others' ignorance of their utilities; it is unethical to convince someone to commit to an action when you know their expected reward is too high and/or their expected costs are too low. By the same token, agents with selective knowledge of their utilities can incur special moral credit or blame: It is particularly admirable to commit to a helpful action when you are ignorant of any extrinsic reward; it is particularly heinous to knowingly perform a costly (e.g., planned and premeditated) harmful action. In short, a wide range of intuitions underlying our judgments of others' competence and values involve considering how agents' might maximize their utilities given subjective and objective elements of costs and reward. In this way, a naïve utility calculus may play a critical role in social evaluation broadly.

Concluding Remarks

The connection between the naïve utility calculus as an account of intuitive decision-making and formal theories of decision-making developed in economics may appear coincidental or simply convenient, but we believe the relation runs deep. As Fritz Heider argued [38], scientific theories, especially in their early stages, may be grounded on commonsense; what better way to formulate initial hypotheses if not by what we intuitively believe to be true? Heider quotes the physicist Robert Oppenheimer: "…all sciences arise as refinement, corrections, and adaptations of common sense."

Suppose that scientific theories of human decision making, starting with classical utility theory and moving through their descendants in behavioral economics, really began grounded on the common sense theory we discussed here. This view has several implications. First, the reason that our models of common-sense psychology in children look like classical utility theory might be because early economists were, with a different purpose in mind, doing exactly what we do here: formalizing common-sense psychology. Second, our common-sense psychology is, at its core, right. Despite the memorable cases where we fail to understand each other, we get others right more often than not. Even if it fails to account for human decision-making in less ecologically relevant domains (e.g., economic choices in the modern marketplace), the naïve utility calculus, as the first models in utility theory, captures key features of human intentional action in the most basic everyday situations even the youngest children appreciate. And as Heider observed, even when commonsense psychology is wrong with respect with how we make choices, it's still right in an important sense. Our most important everyday choices involve others, and our ability to reason about their own choices influences what we do. This intuitive decision-theory is therefore, by definition, a cornerstone of any scientific theory of human decision making.

Finally, the ways in which people's decision making fails to conform with basic assumptions of classical utility theory, which are often counterintuitive and surprising, are surprising precisely because they go against our common-sense. As such, these surprises may point to features of the naïve theory that we have not yet elucidated. To cite just one salient example, we may overinterpret others' failures to help in a low-cost situation as a sign that they don't value helping us. But maybe our naïve theories do not sufficiently take into account agents' non-optimal planning; they wanted to help but they didn't plan well. Or perhaps our naïve theories oversimplify by assuming we know all the relevant costs (or rewards) even when we don't, or assuming that others' costs are like ours even when they're not; both of these assumptions could lead us to mistake a failure to help as a low-cost refusal even when it isn't. Understanding how our commonsense psychology is oversimplified in these ways could advance not only our understanding of core social cognition as scientists, but also, ultimately, help us better understand each other as human beings.

Acknowledgments

This material is based upon work supported by the Center for Brains, Minds, and Machines (CBMM), funded by NSF- STC award CCF-1231216, and by the Simons Center for the Social Brain. We thank Gergely Csibra for useful comments.

Box 1. The Naïve Utility Calculus as a formal computational theory.

For simplicity we assume that costs depend on actions and that rewards depend on states of the world (e.g., being in a desirable physical location or having a specific object), and we focus on deterministic scenarios where agents have perfect information. The formalization can be easily extended to handle more realistic situations.

Generative model

Utility estimation. If *A* is the set of actions that an agent can take (e.g., take a step forwards, pickup an object, etc) and *S* is the set of possible states of the world (determining, for instance, the agent's position in space or her possessions), then a cost function is a mapping $C: A \rightarrow \mathfrak{R}^+$ from actions to costs, and a reward function is a mapping $R: S \rightarrow \mathfrak{R}^+$ from states to rewards.

A plan (or policy) $\pi: S \rightarrow A$ determines what the agent will do in each state in order to arrive at her goal, or final state, s_f, from her initial state s_0. Given cost and reward functions *C* and *R* and a set *P* of plans, a utility function $U_{C,R}: P \rightarrow \mathfrak{R}$ assigns a utility to each plan. In deterministic situations, this utility is the sum of the rewards the agent obtains minus the costs she incurs:

$$U_{C,R}(\pi) = \sum_{S_i=S_o}^{S_f} R(S_i) - C(\pi(S_i)) \tag{1}$$

where s_0 is the starting state, s_f is the target state, s_i are the intermediate states the agent travels through, and $\pi(s_i)$ is the planned action in each of these states.

Plan selection. Because agents' estimates are noisy, they sometimes fail to select the best possible plan. This is modeled through a Boltzmann policy, where the probability of selecting a plan is proportional to

$$p(\pi) \propto \exp\left(\frac{U_{C,R}(\pi)}{\kappa}\right) \tag{2}$$

where $\kappa \in (0, \infty)$ determines the noise in the agent's choice. The smaller the value of κ, the more likely the agent will select high-utility plans.

Inference

Given an agent's actions, the unobservable cost and reward functions can be inferred using Bayes' rule:

$$p(C, R| \text{ Actions}) \propto p(\text{Actions} \mid C, R)\, p(C, R) \tag{3}$$

Here, $\mathrm{p}(\mathrm{C}, \mathrm{R})$ is the prior probability over cost and reward functions, capturing constraints and expectations, and $\mathrm{p}(\text{Actions} \mid \mathrm{C}, \mathrm{R})$ is the likelihood that the agent would take the observed actions given the cost and reward functions. This likelihood term is computed by running the generative model and calculating the probability of the agent selecting each plan and multiplying it by the probability that each plan, in turn, produces the observed actions:

$$\mathrm{p}(\text{Actions} \mid \mathrm{C}, \mathrm{R}) = \int_{\pi \in \mathrm{P}} \mathrm{p}(\text{Actions} \mid \pi)\mathrm{p}(\pi \mid \mathrm{C}, \mathrm{R}) \tag{4}$$

The Bayesian cost and reward inferences specified by Equation 3 are illustrated in Figure I, using an example stimulus from an experiment designed to test the quantitative predictions of this model with adults (who saw a large number of similar stimuli, parametrically varying the agent's path and the configuration of objects and terrain types in the environment).

Box 2: Development of the Naïve Utility Calculus

The studies reviewed here show successes of children in different age groups in different scenarios. Altogether, these open the possibility that some form of the naïve utility calculus is at work from birth. Nevertheless, many aspects of the naïve utility calculus may develop in crucial ways. Here we describe four aspects of the naïve utility calculus that may develop over time.

The dimensions of costs and rewards

As adults we assume that agents' utilities integrate many sources of costs and rewards. Time, effort, attention, or even intangible things like damaging one's reputation can be costly. Similarly, eating, learning, or having a good reputation, for example, can be rewarding. The dimensions that infants consider in utility computation are likely limited and expand over time. For instance, it is not clear how one could assign a cost to breaking social norms without knowing what these social norms are.

Properties of cost and reward functions

As different sources of costs and rewards increase or diminish, so do the costs and rewards. This relation, however, is not linear. For example, the cost associated with exhaustion from walking increases as a function of the distance, but the first steps are less costly than the last ones. Similarly, eating is usually highly rewarding, but eating too little or too much is not. Even if children understand that certain actions or outcomes are costly or rewarding, learning the shape of the cost and reward functions may develop.

Development of the representation of costs and rewards

As adults, we understand that agents act based on their expected costs and rewards. As such, they select the goal with the highest expected utility. In contrast, infants may assume that agents know and act upon exact costs and rewards and over time learn that this is not the case.

Agent-independent priors on costs and rewards

Although individual differences in agents' subjective costs and rewards can only be learned from individuals themselves, agents largely overlap on what they like and dislike. For instance, most people agree that eating sweets is rewarding and that spending time is costly. These priors help observers zoom in on the appropriate cost and reward decompositions. Are these priors learned by finding similarities in costs and rewards across agents? Or do we initially assume all agents have the same costs and rewards, and later infer individual variations across agents?

Figure captions

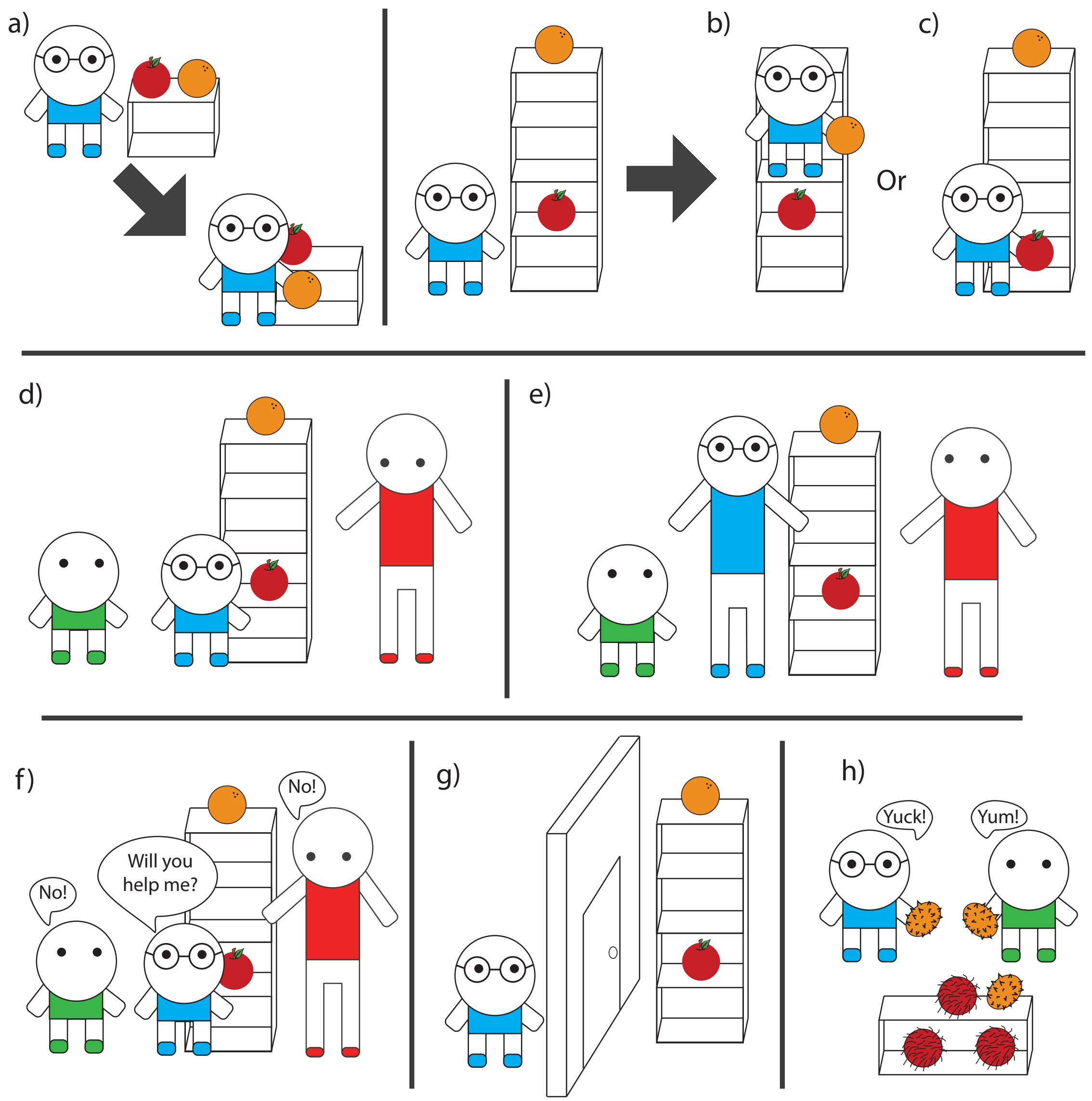

Figure 1 (The logic of costs and rewards underlying out commonsense psychology): (a) If the blue agent wearing glasses (the protagonist) chooses the orange over the apple, how confident are you that she prefers oranges in general to apples? (b) If the orange were high on the top shelf and the agent climbs up to get it, would you become more confident she prefers oranges in general? (c) What if she had chosen the apple instead? Does this indicate any strong preference for apples? (d) If the protagonist wants the orange from the top shelf, whom should she ask for help? (e) If she is the tallest person in the room, is it still appropriate for her to ask for help? (f) If both the red and green agent refuse to help, are they equally mean or is the red one meaner? (g) If the protagonist cannot see the shelf and says she is going to get the orange, are you confident she won't change her mind? (h) If both agents choose kiwanos over rambutans, but one says "yum" and the other says "yuck" after tasting it, who's more likely to have never tasted the fruits before?

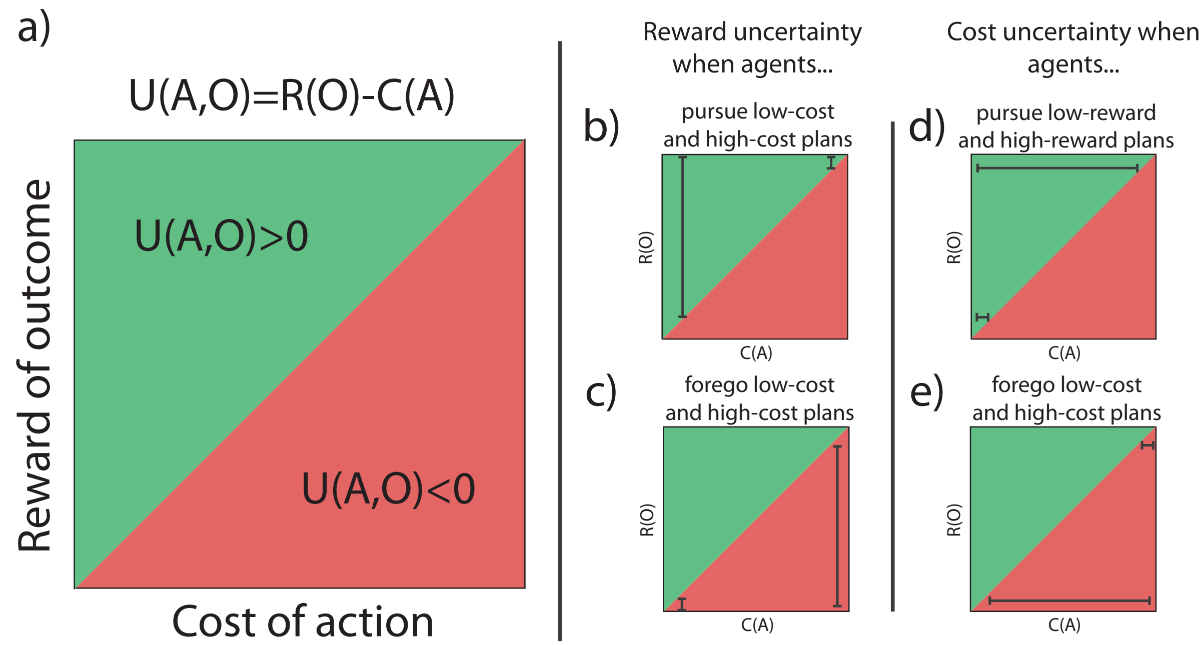

Figure 2 (Key predictions of the naïve utility calculus): Utility maximization in scenarios where costs depend on actions and rewards depend on outcomes. (a) Qualitative plot of how the decision to act reflects the total utility. The utility is positive when the rewards outweigh the costs and negative otherwise. (Green indicates the agent will act and red that she will not.) (b) If an agent pursues a low-cost plan, a wide range of rewards could have produced a positive utility, thus low-cost actions do not reveal much about the agent's reward. However, if the agent pursues a high-cost plan, we can be more confident that the reward was high. (c) The structure of these inferences flips when the agent refuses to pursue a plan. For low cost plans, the rewards must be low to make the net utility negative, thus we can be confident that a decision to forego acting indicates a small reward. By contrast, a wide range of rewards is consistent with refusing to pursue a costly plan, so if the agent refuses to pursue a high cost plan, her refusal is not very informative about her rewards. (d) The implications are parallel when we infer costs given reward knowledge. Low rewards only motivate action when the costs are low, while high rewards motivate action under a wide range of costs; thus pursuit of a goal when rewards are low is more informative about an agent's costs (that they are probably low). (e) If an agent foregoes a low reward we may be uncertain about the costs of acting but if she foregoes a high reward, we can be more certain the costs were high.

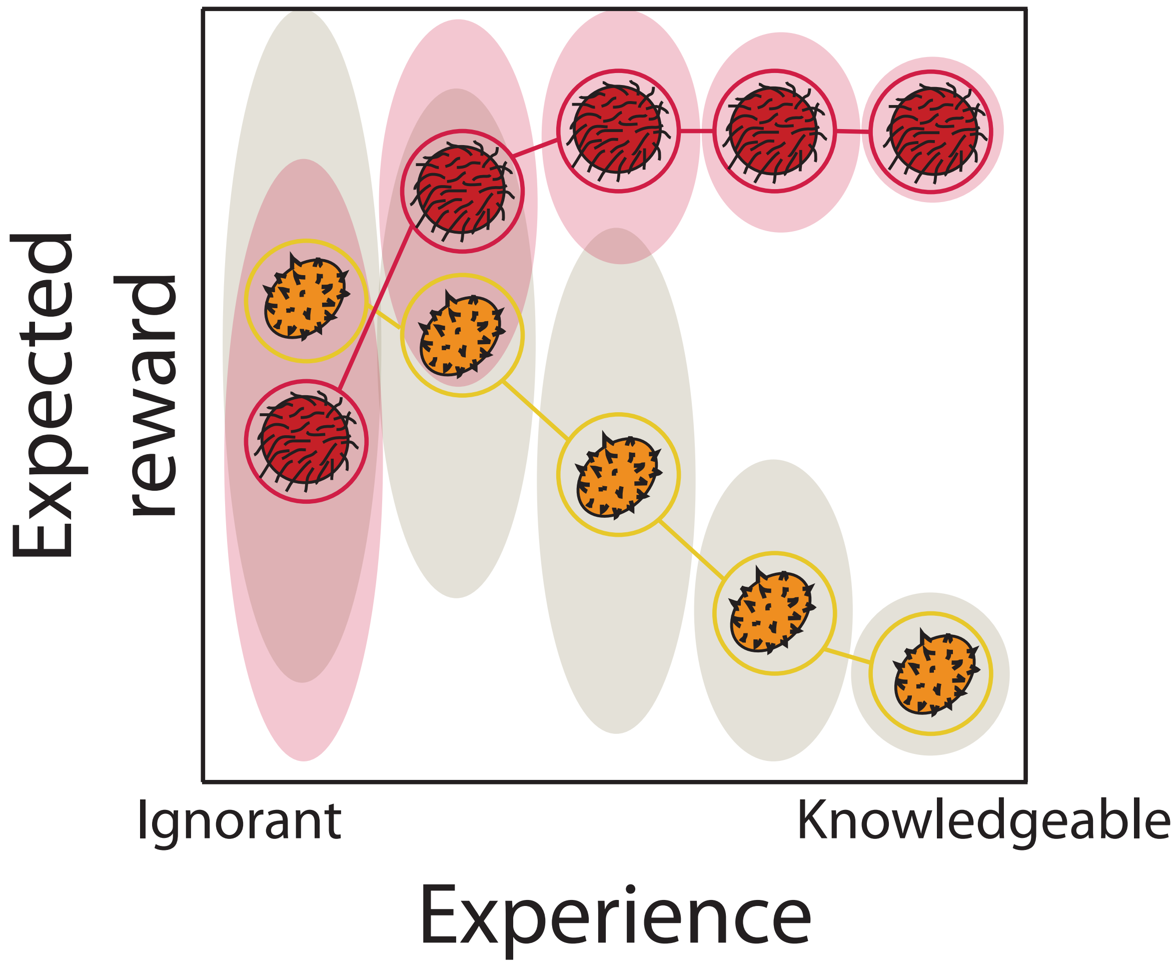

Figure 3 (Relation between experience and reward estimation): Schematic of utility maximization as a function of experience. Ignorant agents (leftmost side of x-axis) have high uncertainty about potential rewards (represented as ovals on the y-axis), and thus are more likely to misidentify the highest-utility option, and more likely to revise their choices over time. Even one encounter with a novel object may cause a large change in the agent's estimate of their expected reward.

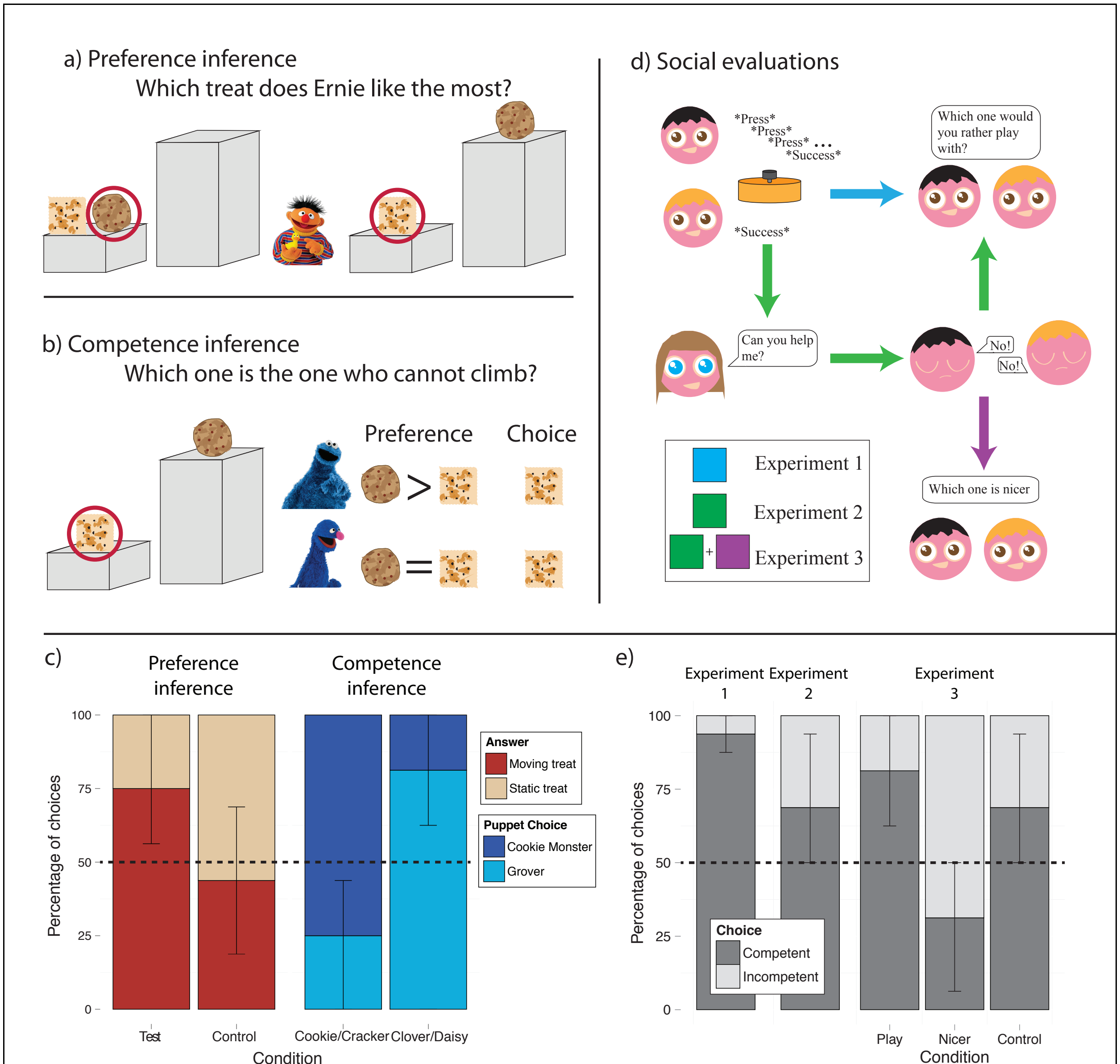

Figure 4 (Empirical support for the naïve utility calculus): Experiments evaluating the Naïve Utility Calculus. (a) Experiment 1 in [57]. Children infer which treat Ernie prefers after watching him pick each one once. (b) Experiment 2 in [57]. Children infer which puppet cannot climb when two agents with different preferences make the same low-cost choice. (c) Results from Experiment 1 (test condition and a control condition) and Experiment 2 (condition where Cookie Monster cannot climb, shown in (b) and a parallel version where Grover cannot climb). (d) Experiments in [59]. (e) Results from the experiments. The control condition in Experiment 3 (not shown in panel (d)) consisted of a baseline measurement of children's judgment of which of the two puppets is nicer.

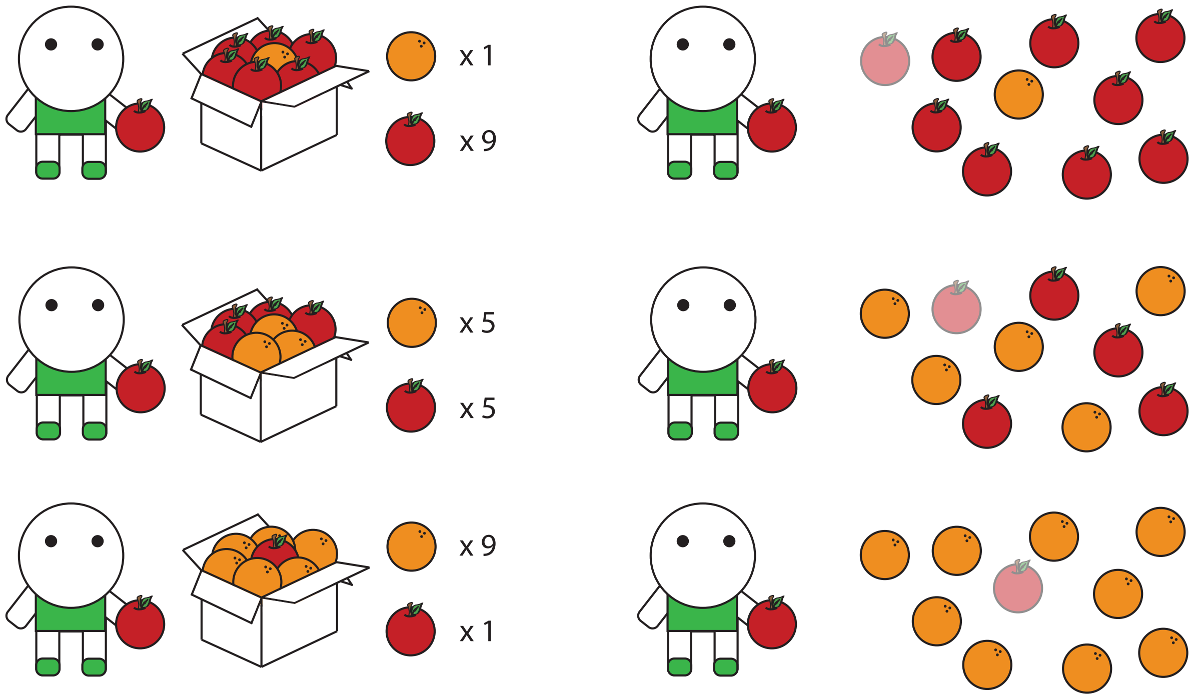

Figure 5 (Costs in statistical and spatial contexts): Graded inferences about preferences in sampling scenarios along with equivalent scenarios unfolded spatially, showing how the assumption of utility maximization supports both types of inferences.

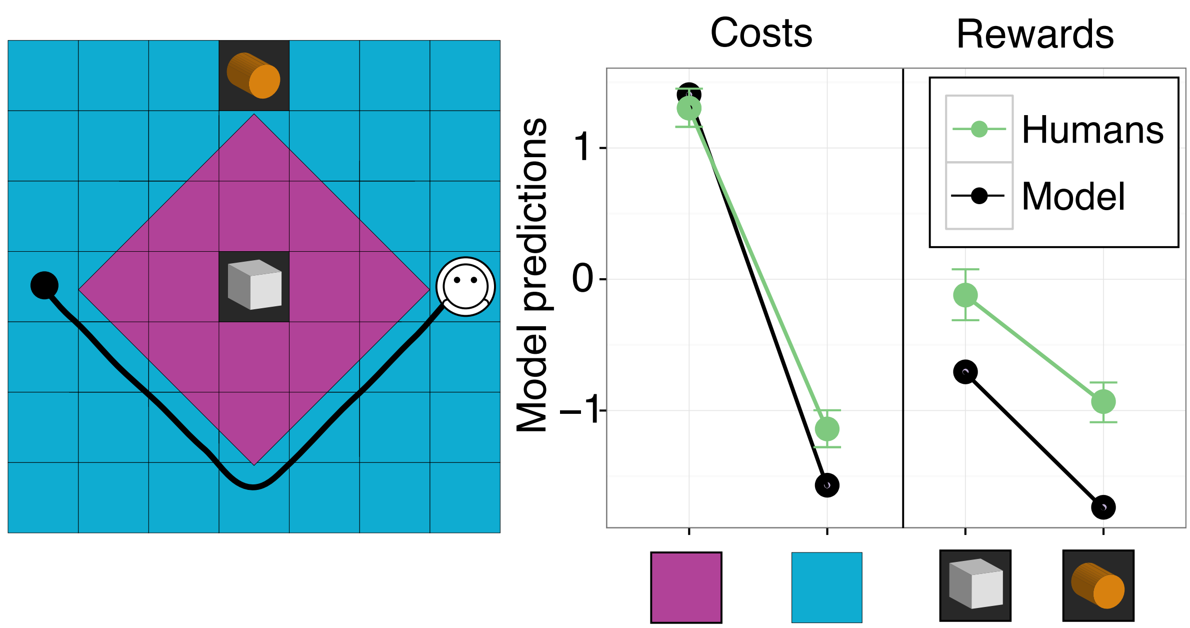

Figure I (Box 1): An agent navigates from a starting point (middle left square) to a target destination (middle right square). If she desires, the agent can also collect the orange package, the white package, or neither. The model infers that the blue terrain is easier to cross than the pink terrain, and that the orange package is less valuable than the white package.

References

1. Biro, S., Verschoor, S., Coalter, E., & Leslie, A. M. (2014). Outcome producing potential influences twelve-month-olds' interpretation of a novel action as goal-directed. *Infant Behavior and Development*, *37*(4), 729-738.
2. Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*(2), 165-193.
3. Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naıve theory of rational action. *Trends in cognitive sciences*, *7*(7), 287-292.
4. Luo, Y., & Baillargeon, R. (2005). Can a self-propelled box have a goal? Psychological reasoning in 5-month-old infants. *Psychological Science*,*16*(8), 601-608.
5. Meltzoff, A. N. (1995). Understanding the intentions of others: re-enactment of intended acts by 18-month-old children. *Developmental psychology*, *31*(5), 838.
6. 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.
7. Woodward, A. L., & Sommerville, J. A. (2000). Twelve-month-old infants interpret action in context. *Psychological Science*, *11*(1), 73-77.
8. Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*(1), 1-34.
9. Kushnir, T., Xu, F., & Wellman, H. M. (2010). Young children use statistical sampling to infer the preferences of other people. *Psychological Science*, *21*(8), 1134-1140.
10. Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: evidence from 14- and 18-month-olds. *Developmental psychology*, *33*(1), 12.
11. Wellman, H. M., Kushnir, T., Xu, F., & Brink, K. A. (2016). Infants Use Statistical Sampling to Understand the Psychological World. *Infancy*.
12. Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in cognitive sciences*, *14*(3), 110-118.
13. Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs?. *Science*, *308*(5719), 255-258.
14. Perner, J. (1991). *Understanding the representational mind*. The MIT Press.
15. Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, *18*(7), 587-592.
16. Wellman, H. M. (1990). The child's theory of mind.
17. Wellman, H. M. (2014). *Making minds: How theory of mind develops*. Oxford University Press.
18. Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child development*, *72*(3), 655-684.
19. Kuhlmeier, V., Wynn, K., & Bloom, P. (2003). Attribution of dispositional states by 12-month-olds. *Psychological Science*, *14*(5), 402-408.
20. Hamann, K., Warneken, F., Greenberg, J. R., & Tomasello, M. (2011). Collaboration encourages equal sharing in children but not in chimpanzees.*Nature*, *476*(7360), 328-331.
21. Hamlin, J. K. (2013). Failed attempts to help and harm: Intention versus outcome in preverbal infants' social evaluations. *Cognition*, *128*(3), 451-474.

22. Hamlin, J. K., Ullman, T., Tenenbaum, J., Goodman, N., & Baker, C. (2013). The mentalistic basis of core social cognition: experiments in preverbal infants and a computational model. *Developmental science*, *16*(2), 209-226.
23. Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, *450*(7169), 557-559.
24. Powell, L. J., & Spelke, E. S. (2013). Preverbal infants expect members of social groups to act alike. *Proceedings of the National Academy of Sciences*,*110*(41), E3965-E3972.
25. Sloane, S., Baillargeon, R., & Premack, D. (2012). Do infants have a sense of fairness?. *Psychological science*, 0956797611422072.
26. Sommerville, J. A., Schmidt, M. F., Yun, J. E., & Burns, M. (2013). The development of fairness expectations and prosocial behavior in the second year of life. *Infancy*, *18*(1), 40-66.
27. Tomasello, M. (2009). *The cultural origins of human cognition*. Harvard University Press.
28. Warneken, F., & Tomasello, M. (2007). Helping and cooperation at 14 months of age. *Infancy*, *11*(3), 271-294.
29. Warneken, F., & Tomasello, M. (2009). Varieties of altruism in children and chimpanzees. *Trends in cognitive sciences*, *13*(9), 397-402.
30. Carey, S. (2009). *The origin of concepts*. Oxford University Press.
31. Meltzoff, A. N. (2007). 'Like me': a foundation for social cognition. *Developmental science*, *10*(1), 126-134.
32. Spelke, E. S. (2000). Core knowledge. *American psychologist*, *55*(11), 1233.
33. Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories*. Mit Press.
34. Gopnik, A., & Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind & Language*, *7*(1-2), 145-171.
35. Gopnik, A., & Wellman, H. M. (1994). The theory theory. In *An earlier version of this chapter was presented at the Society for Research in Child Development Meeting, 1991*. Cambridge University Press.
36. Smith, A. (1759). The theory of moral sentiments, ed. *DD Raphael & AL Macfie. Liberty Fund.(Original work published in 1759.)[ELK]*.
37. Mill, J. S. (1863). Utilitarianism. Collected Works of John Stuart Mill, Vol. 10.*Essays on Ethics, Religion, and Society*.
38. Heider, F. (1958). *The psychology of interpersonal relations*. Psychology Press.
39. Kelley, H. H. (2003). *An atlas of interpersonal situations*. Cambridge University Press.
40. Brown, R. (1985). Social Psychology. Free Press.
41. Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition*, *113*(3), 329-349.
42. Goodman, N. D., Tenenbaum, J. B. & Gerstenberg, T. (2015). Concepts in a probabilistic language of thought. In *The Conceptual Mind: New Directions in the Study of Concepts*,
43. Gopnik, A., & Wellman, H. M. (2011). Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory theory. *Psychological bulletin*, *138*(6), 1085.
44. Griffiths, T. L., Kemp, C., & Tenenbaum, J. B. (2008). Bayesian models of cognition.

45. Lucas, C. G., Griffiths, T. L., Xu, F., Fawcett, C., Gopnik, A., Kushnir, T., ... & Hu, J. (2014). The child as econometrician: A rational model of preference understanding in children. *PloS one*, *9*(3), e92160.
46. Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *science*, *331*(6022), 1279-1285.
47. Von Neumann, J., & Morgenstern, O. (1953). *Theory of games and economic behavior*. Princeton university press.
48. Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine. *Econometrica: Journal of the Econometric Society*, 503-546.
49. Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
50. Russell, B., (1925). *ABC of Relativiy.* Routledge.
51. Battaglia, P. W., Hamrick, J. B., & Tenenbaum, J. B. (2013). Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences*, *110*(45), 18327-18332.
52. Kelley, H. H. (1973). The processes of causal attribution. *American psychologist*, *28*(2), 107.
53. Ullman, T., Baker, C., Macindoe, O., Evans, O., Goodman, N., & Tenenbaum, J. B. (2010). Help or hinder: Bayesian models of social goal inference. In *Advances in neural information processing systems* (pp. 1874-1882).
54. Markman, E. M. (1991). The whole-object, taxonomic, and mutual exclusivity assumptions as initial constraints on word meanings. *Perspectives on language and thought: Interrelations in development*, 72-106.
55. Griffiths, T. L., Chater, N., Kemp, C., Perfors, A., & Tenenbaum, J. B. (2010). Probabilistic models of cognition: Exploring representations and inductive biases. *Trends in cognitive sciences*, *14*(8), 357-364.
56. Schachner, A., & Carey, S. (2013). Reasoning about 'irrational' actions: When intentional movements cannot be explained, the movements themselves are seen as the goal. *Cognition*, *129*(2), 309-327.
57. Jara-Ettinger, J., Gweon, H., Tenenbaum, J. B., & Schulz, L. E. (2015). Children's understanding of the costs and rewards underlying rational action. *Cognition*, *140*, 14-23.
58. Jara-Ettinger, J., Lydic, E., Tenenbaum, J. B., & Schulz, L. E. Beliefs about desires: Children's understanding of how knowledge and preference influence choice.
59. Jara-Ettinger, J., Tenenbaum, J. B., & Schulz, L. E. (2015). Not So Innocent Toddlers' Inferences About Costs and Culpability. *Psychological science*, *26*(5), 633-640.
60. 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.
61. Csibra, G. (2008). Action mirroring and action understanding: An alternative account. *Sensorymotor Foundations of Higher Cognition. Attention and Performance XXII*, 435-459.
62. Scott, R. M., & Baillargeon, R. (2013). Do infants really expect agents to act efficiently? A critical test of the rationality principle. *Psychological science*, 0956797612457395.

63. 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.
64. Southgate, V., Johnson, M. H., & Csibra, G. (2008). Infants attribute goals to even biologically impossible actions. Cognition, 107, 1059-1069.
65. Csibra, G., Gergely, G., Bíró, S., Koos, O., & Brockbank, M. (1999). Goal attribution without agency cues: the perception of 'pure reason'in infancy. *Cognition*, *72*(3), 237-267.
66. Johnson, S. G., & Rips, L. J. (2015). Do the right thing: The assumption of optimality in lay decision theory and causal judgment. *Cognitive psychology*,*77*, 42-76.
67. Gergely, G., Bekkering, H., & Király, I. (2002). Rational imitation in preverbal infants. *Nature*, *415*(6873), 755-755.
68. Premack, D. (1990). The infant's theory of self-propelled objects. *Cognition*,*36*(1), 1-16.
69. Saxe, R., Tenenbaum, J. B., & Carey, S. (2005). Secret agents inferences about hidden causes by 10-and 12-month-old infants. *Psychological Science*,*16*(12), 995-1001.
70. Leslie, A. M. (1984). Infant perception of a manual pick-up event. *British Journal of Developmental Psychology*, *2*(1), 19-32.
71. Muentener, P., & Carey, S. (2010). Infants' causal representations of state change events. *Cognitive Psychology*, *61*(2), 63-86.
72. Saxe, R., Tzelnic, T., & Carey, S. (2007). Knowing who dunnit: Infants identify the causal agent in an unseen causal interaction. *Developmental Psychology*,*43*(1), 149.
73. Newman, G. E., Keil, F. C., Kuhlmeier, V. A., & Wynn, K. (2010). Early understandings of the link between agents and order. *Proceedings of the National Academy of Sciences*, *107*(40), 17140-17145.
74. Ma, L. & Xu, F. (2013) Preverbal Infants Infer Intentional Agents From the Perception of Regularity. *Developmental Psychology, (48, 7)*, 1330-1337.
75. Muentener, P., & Schulz, L. (2014). Toddlers infer unobserved causes for spontaneous events. *Frontiers in psychology*, *5*.
76. Wu, Y., Muentener, P., & Schulz, L. E. (2015). The Invisible Hand: Toddlers Connect Probabilistic Events With Agentive Causes. *Cognitive science*.
77. Skerry, A. E., & Spelke, E. S. (2014). Preverbal infants identify emotional reactions that are incongruent with goal outcomes. *Cognition*, *130*(2), 204-216.
78. Denison, S., Reed, C., & Xu, F. (2013). The emergence of probabilistic reasoning in very young infants: Evidence from 4.5-and 6-month-olds.*Developmental Psychology*, *49*(2), 243.
79. Denison, S., & Xu, F. (2010). Integrating Physical Constraints in Statistical Inference by 11-Month-Old Infants. *Cognitive Science*, *34*(5), 885-908.
80. Dewar, K. M., & Xu, F. (2010). Induction, overhypothesis, and the origin of abstract knowledge evidence from 9-month-old infants. *Psychological Science*.
81. Xu, F., & Denison, S. (2009). Statistical inference and sensitivity to sampling in 11-month-old infants. *Cognition*, *112*(1), 97-104.
82. Xu, F., & Garcia, V. (2008). Intuitive statistics by 8-month-old infants. *Proceedings of the National Academy of Sciences*, *105*(13), 5012-5015.

83. Gweon, H., Tenenbaum, J. B., & Schulz, L. E. (2010). Infants consider both the sample and the sampling process in inductive generalization. *Proceedings of the National Academy of Sciences*, *107*(20), 9066-9071.
84. Xu, F., & Tenenbaum, J. B. (2007). Sensitivity to sampling in Bayesian word learning. *Developmental science*, *10*(3), 288-297.
85. Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological review*, *114*(2), 245.
86. Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., & Schulz, L. (2011). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition*, *120*(3), 322-330.
87. Gweon, H., Pelton, H., Konopka, J. A., & Schulz, L. E. (2014). Sins of omission: children selectively explore when teachers are under-informative.*Cognition*, *132*(3), 335-341.
88. Gweon, H., Shafto, P., & Schulz, L. E. (2014). Children consider prior knowledge and the cost of information both in learning from and teaching others. In *Proceedings of the 36th annual conference of the Cognitive Science Society*.
89. Liszkowski, U., Carpenter, M., & Tomasello, M. (2008). Twelve-month-olds communicate helpfully and appropriately for knowledgeable and ignorant partners. *Cognition*, *108*(3), 732-739.
90. Matthews, D., Lieven, E., Theakston, A., & Tomasello, M. (2006). The effect of perceptual availability and prior discourse on young children's use of referring expressions. *Applied Psycholinguistics*, *27*(3), 403.
91. Gweon, H., Chu, V., & Schulz, L. E. (2014). To give a fish or to teach how to fish? Children weigh costs and benefits in considering what information to transmit. In *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 559-564). Cognitive Science Society Austin, TX.
92. Bridgers, S., Jara-Ettinger, J., & Gweon, H (2016). Children consider others' expected costs and rewards when deciding what to teach. In *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 559-564). Cognitive Science Society Austin, TX.
93. Ronfard, S., Was, A. M., & Harris, P. L. (2016). Children teach methods they could not discover for themselves. *Journal of experimental child psychology*,*142*, 107-117.
94. Grice, H. P. (1970). *Logic and conversation*.
95. Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, *336*(6084), 998-998.
96. Stiller, A. J., Goodman, N. D., & Frank, M. C. (2015). Ad-hoc implicature in preschool children. *Language Learning and Development*, *11*(2), 176-190.
97. Hamlin, J. K., Wynn, K., Bloom, P., & Mahajan, N. (2011). How infants and toddlers react to antisocial others. *Proceedings of the national academy of sciences*, *108*(50), 19931-19936.
98. Hamlin, J.K. (2014). Context-dependent social evaluation in 4.5-month-old human infants: The role of domain-general versus domain-specific processes in the development of social evaluation. Frontiers in Psychology.