CellPress

Feature Review

The Naïve Utility Calculus: Computational Principles Underlying Commonsense Psychology

Julian Jara-Ettinger,[1,*] Hyowon Gweon,[2] Laura E. Schulz,[1] and Joshua B. Tenenbaum[1,*]

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' allows both children and adults observe the behavior of others 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 inception of the field, 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 the goals [1–8], desires [9–11], beliefs [12–18], and pro-social behavior [19–29] of other agents has advanced our understanding of what in our commonsense psychology is at work in early infancy [30–32] and what develops later [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?

Trends

A growing number of studies have found that people have a commonsense theory of psychology – a cognitive framework for making inferences about the goals of others and their preferences, competencies and motivations, experiences, and beliefs – that is already at work in surprisingly sophisticated ways from early childhood.

In toddlers and children, commonsense psychology appears to be guided by the assumption that agents maximize utilities, or tradeoffs between rewards and costs of action.

Computational models that capture quantitative aspects of the social inferences of adults are also guided by an expectation of utility maximization, embedded in a Bayesian framework.

Taken together, empirical and computational studies in adults and children converge on the idea that a 'naïve utility calculus' is at the heart of human social cognition.

[1]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[2]Department of Psychology, Stanford University, Stanford, CA 94305, USA

*Correspondence: jjara@mit.edu (J. Jara-Ettinger) and jbt@mit.edu (J.B. Tenenbaum).

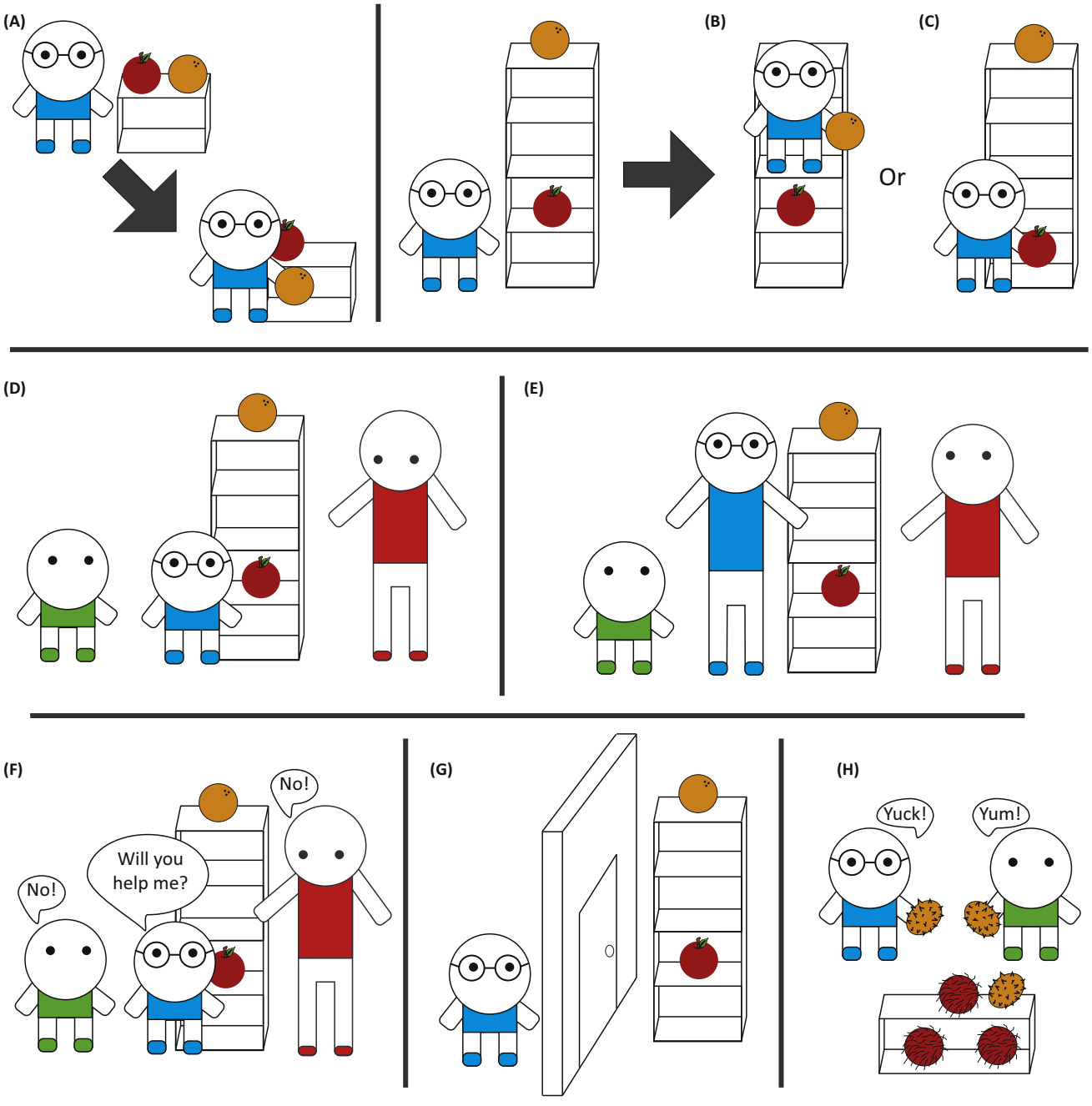

Figure 1. The Logic of Costs and Rewards Underlying Commonsense Psychology. (A) If the blue agent wearing glasses (the protagonist) chooses the orange over the apple, how confident are you that she prefers oranges in general to apples? (B) If the orange were high on the top shelf and the agent climbs 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 is more likely to have never tasted the fruits before?.