Cognition 239 (2023) 105524

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

People can use the placement of objects to infer communicative goals

Michael Lopez-Brau *, Julian Jara-Ettinger

Department of Psychology, Yale University, 2 Hillhouse Ave, New Haven, CT 06520, USA

A R T I C L E I N F O

Dataset link: https://osf.io/57n4g

Keywords:
Computational modeling
Social objects
Theory of Mind

A B S T R A C T

Beyond words and gestures, people have a remarkable capacity to communicate indirectly through everyday objects: A hat on a chair can mean it is occupied, rope hanging across an entrance can mean we should not cross, and objects placed in a closed box can imply they are not ours to take. How do people generate and interpret the communicative meaning of objects? We hypothesized that this capacity is supported by social goal inference, where observers recover what social goal explains an object being placed in a particular location. To test this idea, we study a category of common ad-hoc communicative objects where a small cost is used to signal avoidance. Using computational modeling, we first show that goal inference from indirect physical evidence can give rise to the ability to use object placement to communicate. We then show that people from the U.S. and the Tsimane'—a farming-foraging group native to the Bolivian Amazon—can infer the communicative meaning of object placement in the absence of a pre-existing convention, and that people's inferences are quantitatively predicted by our model. Finally, we show evidence that people can store and retrieve this meaning for use in subsequent encounters, revealing a potential mechanism for how ad-hoc communicative objects become quickly conventionalized. Our model helps shed light on how humans use their ability to interpret other people's behavior to embed social meaning into the physical world.

1. Introduction

Humans have a remarkable capacity to communicate through objects, even ones we do not usually think of as conveying meaning. A hat on a chair can reveal that the seat is taken; rope surrounding a patch of grass can tell us not to walk through; and, during snowy winters in the northeastern United States, plastic chairs on shoveled parking spots are used to signal that they are not up for grabs. These kinds of everyday objects (Fig. 1) do little to physically constrain our actions, yet they affect our behavior because we recognize the meaning they convey. Consistent with this, past empirical research has shown that people spontaneously use objects to communicate (e.g., leaving an open notebook on a library table to mark that the space is occupied; Becker & Mayo, 1971; Edney & Jordan-Edney, 1974; Sommer & Becker, 1969), and detect when an object is communicative (e.g., realizing that the table with a notebook must be taken; Becker, 1973; Shaffer & Sadowski, 1975), with this ability possibly emerging in childhood (Rossano et al., 2015; Zhao et al., 2020).

What are the cognitive capacities that support our ability to communicate through objects? One possibility is that communicative objects emerge from a system of simple conventions, where objects and their placement are explicitly associated with different communicative meanings. As children, for instance, most of us likely ignored strap barriers at banks, movie theaters, and DMVs, and their meaning had to be explicitly taught to us. After learning their meaning, we were then able to recall it whenever we encountered them in new locations.

While conventional knowledge is undoubtedly a major driver for how we learn and use communicative objects, people are also able to generate novel communicative objects that others can readily understand (such as placing an ironing board to mark that someone has reserved a parking spot; Fig. 1i). What computations underlie this capacity? And how does the communicative meaning of novel objects become conventionalized?

Here we hypothesized that the capacity to embed and infer communicative meaning from novel objects emerges from our ability to reason about the mental states behind other people's behavior—our *Theory of Mind* (ToM; Wellman, 2014; Gopnik et al., 1997). The central idea in our proposal is that, if people can infer other agents' mental states based on how they manipulated an object (via Theory of Mind), then people can also strategically manipulate objects with the purpose of eliciting mental-state inferences in agents who encounter these objects. Through this method, people can intentionally manipulate their environment with the goal of communicating their desires to people who navigate the environment when the communicator is absent. We propose that this type of reasoning might support the creation of novel ad-hoc communicative objects, which can then quickly become conventionalized and widespread, supported via memory and recognition.

* Corresponding author.
E-mail addresses: michael.lopez-brau@yale.edu (M. Lopez-Brau), julian.jara-ettinger@yale.edu (J. Jara-Ettinger).

https://doi.org/10.1016/j.cognition.2023.105524
Received 8 September 2022; Received in revised form 9 June 2023; Accepted 10 June 2023
Available online 12 July 2023

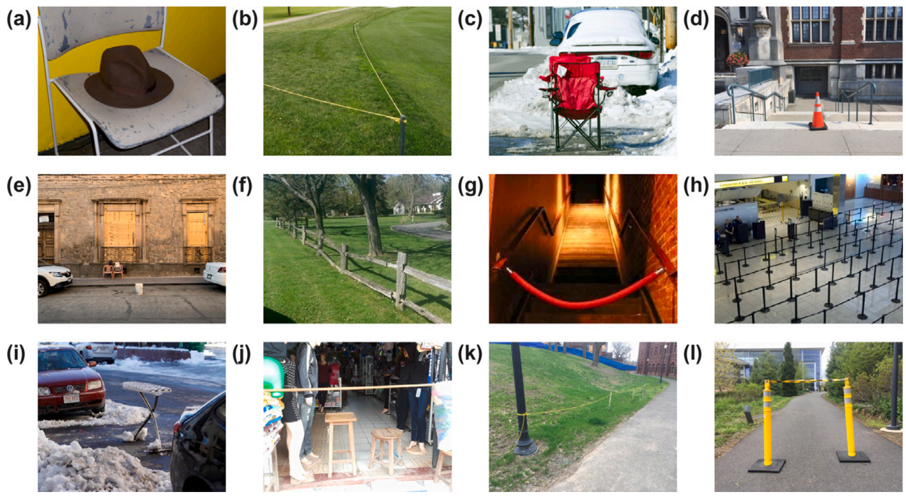

Fig. 1. Real-world examples of people communicating through objects. (a) A hat on a chair indicating that someone intends to return. (b) Rope a few inches above the grass so that people know not to walk through. (c) Chairs along the side of the street in South Boston to reveal that someone shoveled and claimed this parking spot. (d) A traffic cone in front of some stairs signaling limited access. (e) A bucket along the side of the street in central Mexico indicating that the parking spot is reserved. (f) An easy-to-cross fence marking a property limit. (g) A stanchion across a stairwell revealing access may be restricted to certain individuals. (h) Belt barriers at the airport telling passengers that they should form a line (and where). (i) An ironing board along the side of the street indicating that the parking spot is taken. (j) A wooden pole and two small benches in a store in Bolivia indicating that the owner is not available. (k) A small rope along a sidewalk asking people not to walk near a construction site. (l) A pair of traffic posts deterring people from using this walkway.

To explore this idea, this paper focuses on a family of objects like those shown in Fig. 1. These objects are often not intrinsically communicative: Hats, chairs, and rope are not purposefully designed for communication, but they can nonetheless convey a message when placed in certain locations (Figs. 1a–c). Moreover, despite their varied use, these objects all communicate some kind of restriction (e.g., "do not use" or "do not cross"). Critically, however, this restriction is not imposed purely through a physical constraint: The cost that these objects impose on agents is low enough that it could be easily ignored (e.g., walking over the "barriers" in Figs. 1b and 1k is trivial). Intuitively, these objects instead work because people realize that the object was intentionally placed with the purpose to communicate. Because of this common structure, we will refer to these objects as *low-cost communicative blockers* (LCCBs). While these objects do not capture the full scope of everyday communicative objects, we believe their use is a fruitful case study for understanding our proposal. We return our focus to communicative objects more broadly in the Discussion.

To illustrate the logic of our proposal, imagine trying to find the exit of an unfamiliar building. As you walk down a hallway, you find two doors, side by side. Suppose, however, that one of the doors has a broom positioned diagonally across it. Naturally, it is easy to recognize that (1) someone intentionally placed the broom there and that (2) it creates a small inconvenience for people wanting to walk through the door. When considering why someone would choose to use a broom to block a door, one possibility is that they wanted to prevent people from walking through. But if that were the case, why not put more effort into blocking the door, given how easy it is to move the broom out of the way? Intuitively, this is because their goal was not to create an insurmountable physical constraint—which would require more effort to achieve—but rather to prompt you to infer that they do not want you to walk through.

This proposal assumes that people can detect intentional arrangements of objects (e.g., a broom placed diagonally across a door was likely placed intentionally), infer what an agent did (e.g., an agent must have taken the broom and placed it there), and determine how much effort it required from the agent and how much it affects us (e.g., how hard was it to place the broom and what effects does this have on my potential plans?). Consistent with this, past research has shown that people have a rich understanding of what physical environments reveal about people (Gosling et al., 2002; Hurwitz et al., 2019). Moreover, people can infer others' actions from indirect physical evidence of their presence (Lopez-Brau et al., 2022), and estimate the effort involved in moving and manipulating objects (Yildirim et al., 2019). These capacities also emerge early in development, with children drawing surprisingly rich inferences from physical evidence, ranging from inferences about what actions an agent took (Jacobs et al., 2021) and what they knew (Pelz et al., 2020) to inferences about even richer social information, such as whether two people transmitted ideas (Pesowski et al., 2020) and have shared interests (Pesowski et al., 2021).

Critically, for communicative objects to have their intended effect, the ability to reason about them is not enough: people must also be motivated to behave cooperatively. If this were not the case, people would ignore low-cost communicative blockers (LCCBs; since the cost they impose is negligible), and communicators would favor creating insurmountable physical constraints rather than communicative signals. While there are undoubtedly cases where people ignore LCCBs, and where people build physical barriers because they do not expect cooperativeness, the pervasive use of these objects suggests that there are many cases where people expect strangers to cooperate by default. This is consistent with evidence that even young children will spontaneously cooperate with strangers (Warneken & Tomasello, 2006) and that adults have a default propensity to cooperate (Rand, 2016).

While all this past work establishes the cognitive pre-requisites that our proposal builds on, to our knowledge, no work has yet explored specifically whether these capacities underlie the ability to use communicative objects (Fig. 1).

1.1. Paper overview

Our paper has three goals. Our first goal is to test whether our theoretical proposal can, in principle, explain the logic of low-cost communicative blockers (LCCBs), where agents share mental states by using objects to impose a minimal cost on observers. To achieve this, we present a model that explicitly formalizes our proposal in computational terms, and we explore its behavior in synthetic simulations (Sections 2–3) to test whether it can produce patterns that resemble how people use communicative objects. Our computational model focuses on the inferences that we make once an object placement is detected as intentional. We return to the question of how to detect intentional placements in the Discussion.

Having found theoretical support for our proposal, our second goal is to test whether the fine-grained quantitative predictions of our account match human judgments when reasoning about novel low-cost communicative blockers. That is, we use our model to generate exact numerical predictions about the strength of inferences that people should make in different situations, and we compare them to human judgments (Experiment 1).

Finally, our third goal is to test whether the mechanisms we propose play a role when conventional knowledge (i.e., object-meaning mappings that are in common ground for a social group) is unavailable. That is, our goal is not to argue against the critical role of convention, but rather to ask what types of inferences people engage in when they face an object that has no conventional meaning attached to it (Experiments 2–3), and to explore how these inferences become conventionalized (Experiment 4).

While our account proposes mechanisms that support both the creation and understanding of communicative objects, these two behaviors are asymmetrical in two ways. First, as we show in our model below, recognizing the meaning of communicative objects is easier than creating them, requiring one fewer level of recursion. Second, for communicative objects to become ubiquitous, the ability to infer their meaning must be widespread, while the ability to invent them can be restricted to a few individuals. Thus, after confirming the computational plausibility of our account (Sections 2–3), our behavioral studies (Experiments 1–4) focus on people's ability to infer the communicative meaning of low-cost communicative blockers, rather than on how they are created. We return to this asymmetry in cognitive demand in the Discussion.

2. Computational framework

For simplicity and clarity, we describe our model in the context of a simple event similar to the ones we use in Experiments 2–4. Here, an agent (the *decider*) encounters two doors—door *A* and door *B*—and must decide which one to walk through. Before they do, another agent (the *enforcer*), who wants to influence the decider's choice, has the opportunity to place objects in front of either door, including stacking multiple objects to create a physical constraint. To illustrate, Fig. 2a shows a situation where the enforcer has access to four boulders that can be positioned in front of either door. Figs. 2d–h show five possible changes that the enforcer could implement (among many others). Our model therefore consists of (i) an enforcer that moves objects in a scene with the goal of affecting a decider's behavior, and (ii) a decider that determines what to do in a scene by thinking about the costs that the objects in the scene impose. Critically, we assume that the enforcer and decider are never in the scene at the same time, such that the decider only has access to the physical layout of the scene.

In this setup, the enforcer can always pursue a simple non-communicative strategy: stack enough objects in front of one of the doors to the point that walking through it is so much work that the decider will prefer to avoid it (e.g., Fig. 2f). However, stacking objects in front of a door is also costly for the enforcer. This creates a preference for more efficient strategies, where agents might exploit their Theory of Mind to use objects in a communicative manner.

Under our proposal, people use objects to share mental states by reasoning about the costs incurred by the enforcer (how costly is it for the enforcer to block paths?) and the decider (what costs does this impose on the decider?). Because past work has already studied how people reconstruct behavior from physical displays and estimate the underlying costs (Lopez-Brau et al., 2022; Pesowski et al., 2020; Yildirim et al., 2019), our model takes this capacity for granted and focuses on the inferences that people make given access to these costs.

To make cost-based inferences, our framework instantiates Theory of Mind (ToM) as a form of simple recursive social reasoning, similar to models developed to understand pedagogical demonstrations (Ho et al., 2016; Shafto et al., 2014), pragmatics (Frank & Goodman, 2012; Goodman & Frank, 2016), and mental-state inferences (Ullman et al., 2009), and similar to the logic behind *k*-level ToM models (where *k* is a variable indicating the recursion depth within ToM; Devaine et al., 2014). At its core, our model is structured around an assumption that agents act to maximize their subjective utilities—the difference between the costs that they incur and the rewards that they obtain. This assumption is at the heart of human mental-state inferences in adults (Baker et al., 2017; Jara-Ettinger et al., 2020; Jern et al., 2017) and emerges early in development (Gergely & Csibra, 2003; Jara-Ettinger et al., 2016; Liu et al., 2017; Lucas et al., 2014).

Formally, let S be the space of all possible scenes, where each scene $s \in S$ represents an observable arrangement of objects (e.g., see Figs. 2a, d–h for six possible scenes in the boulder example). Each agent in this context is defined by two main components. The first is a cost function that captures how agents interact with the environment. For the enforcer, their cost function C_E represents the cost of moving objects, such that $C_E(s_0, s)$ is the cost of transforming an initial scene s_0 into a final scene s (e.g., the cost of changing the scene so that an object in a corner is now in front of a door). For the decider, their cost function C_D represents the cost of navigating the environment, such that $C_D(a, s)$ is the cost of taking action a in scene s (e.g., the cost of walking through a door with an object in the way).

The second main component is a reward function that captures each agent's desires. The enforcer's reward function R_E represents their desire to affect the decider's behavior. That is, $R_E(a)$ is the reward the enforcer obtains when the decider takes action a. The decider's reward function R_D represents their own personal desires: $R_D(a)$ is the decider's personal reward when choosing action a (e.g., the decider's reward when they choose door A in Fig. 2).

Our computational framework uses these cost and reward functions to build a model of recursive social reasoning, where the enforcer decides how to move objects by thinking about what action they hope the decider will take, and the decider decides what action to take by inferring what the enforcer wants, based on how they manipulated the objects in the scene. Below, we present the logic of our model, starting with the grounding level of the recursive structure. A more detailed presentation of our model can be found in SM.

2.1. Non-mentalistic decider

The lowest level of our model consists of a non-mentalistic decider D_0 that represents an agent lacking any awareness that objects in a scene may have been intentionally manipulated by another agent. This decider therefore chooses what to do based on the physical properties of the scene alone. Given a scene s, the non-mentalistic decider's utility for taking action a is given by:

$$U_{D_0}(a; s) = R_{D_0}(a) - C_{D_0}(a, s), \quad (1)$$

where $R_{D_0}(a)$ is the reward that the decider obtains from taking action a and $C_{D_0}(a, s)$ is the cost they incur from taking that action in scene s.

We transform this utility function into a probability distribution over actions by applying the *softmax function*:

$$p_{D_0}(a|s) \propto \exp(U_{D_0}(a; s)/\tau). \quad (2)$$

The softmax function is a standard method for transforming utility functions into probability distributions, guided by a temperature parameter $\tau \in (0, \infty)$. When τ is low, the decider consistently chooses the actions that maximize the utility function (converging towards optimal behavior as $\tau \to 0$). When τ is high, the decider's behavior becomes noisier, and the agent is more likely to select actions that are not necessarily the best ones (converging towards random behavior as $\tau \to \infty$).

2.2. Simple enforcer

The next level of our model consists of a simple enforcer E_0 who reasons about the non-mentalistic decider D_0. That is, this enforcer determines what to do under the assumption that the decider will not realize that the objects contain any social information and will instead see them as nothing more than physical obstacles.

Formally, suppose the world is in some initial state s_0 and the enforcer wants the decider to take action a (e.g., the initial scene might be Fig. 2a, and the enforcer wants the decider to choose door A). To do this, the simple enforcer considers different possible scenes s (e.g., Figs. 2d–h) and evaluates them through the utility function:

$$U_{E_0}(s; a, s_0) = \underbrace{R_{E_0}(a) p_{D_0}(a|s)}_{\substack{\text{Expected reward} \\ \text{when decider takes action } a \text{ in scene } s}} - \overbrace{C_{E_0}(s_0, s).}^{\substack{\text{Cost of transforming} \\ \text{scene } s_0 \text{ into scene } s}} \tag{3}$$

Here, the first term ($R_{E_0}(a) p_{D_0}(a|s)$) is the enforcer's expected reward (i.e., the reward they obtain when the decider takes action a, weighted by the probability that the decider takes this action). This enforcer's ability to predict how the decider will act, $p_{D_0}(a|s)$, is computed using the non-mentalistic decider model (Eq. (2), Section 2.1). This term is then balanced against the cost $C_{E_0}(s_0, s)$ that the enforcer incurs in transforming scene s_0 into scene s. Combined, the first term leads the enforcer to prefer scenes where the decider is more likely to take the desired action, and the second term leads the enforcer to favor minimal scene changes over drastic ones.

2.3. Mentalistic decider

Having defined the simple enforcer E_0, we can now specify a mentalistic decider D_1 that reasons about this enforcer's choices. That is, this decider infers why the enforcer decided to modify the scene, and takes this into account when deciding what to do.

Formally, the mentalistic decider assigns a utility to each action via

$$U_{D_1}(a; s_0, s, \phi) = \underbrace{R_{D_1}(a) - C_{D_1}(a, s)}_{\substack{\text{Decider's egocentric} \\ \text{costs and rewards}}} + \overbrace{\phi \langle \ell(a|s_0, s) \rangle}^{\substack{\text{Decider's allocentric} \\ \text{preferences}}} . \tag{4}$$

The first two terms ($R_{D_1}(a)$ and $C_{D_1}(a, s)$) capture the decider's egocentric rewards and costs for taking action a in scene s, respectively (identical to the utility function for the non-mentalistic decider; Eq. (1), Section 2.1).

The final term, $\phi \langle \ell(a|s_0, s) \rangle$, represents the decider's utility for acting in accordance with the enforcer's preferences, also known as the decider's "adopted utility" (Powell, 2022). Here, $\langle \ell(a|s_0, s) \rangle$ is the decider's belief that the enforcer wants them to take action a, based on the change from scene s_0 to scene s. This term is then weighted by a real-valued cooperation parameter ϕ that captures the decider's motivation to pursue, act against, or ignore the enforcer's preferences. When ϕ is positive, the decider is motivated to act in a way that is consistent with the enforcer's preferences. Conversely, when ϕ is negative, the decider is antagonistic and prefers to act against the enforcer's preferences. Finally, when $\phi = 0$, the decider acts egocentrically (becoming the same model as the non-mentalistic decider), and treats objects as physical constraints, ignoring why the enforcer might have positioned them there.

Critically, the decider does not know a priori what the enforcer wants them to do (i.e., the decider does not have direct access to $\ell(a|s_0, s)$). This term is therefore inferred by considering the enforcer's possible reward functions:

$$\langle \ell(a|s_0, s) \rangle = \int_{R_{E_0} \in R} \mathbb{1}_{\arg\max R_{E_0}}(a) p(R_{E_0}|s_0, s). \tag{5}$$

This equation adds up the probability of every possible reward function $R_{E_0} \in R$ where a is the preferred action. For each of these reward functions, its probability, $p(R_{E_0}|s_0, s)$, is inferred by reasoning about the enforcer's choice to change scene s_0 into scene s:

$$p(R_{E_0}|s_0, s) \propto p_{E_0}(s|R_{E_0}, s_0) p(R_{E_0}), \tag{6}$$

with the likelihood $p_{E_0}(s|R_{E_0}, s_0)$ computed using the simple enforcer model (i.e., it is given by the softmax of Eq. (3), Section 2.2).

Note that this formulation assumes that the mentalistic decider knows both the scene's initial and final states (s_0 and s). This allows deciders to infer the enforcer's rewards by reasoning about the costs that were introduced. In more realistic situations, it is more likely that deciders have a prior distribution over scenes ($p(s_0)$) rather than perfect knowledge about the initial scene s_0. Returning to the example in the introduction, for instance, when encountering a broom placed across an entrance, a decider may not know where the broom was situated before an enforcer placed it across the door, but they may believe it was more likely that it was positioned elsewhere. Modeling prior expectations about scene distributions is beyond the scope of our model, but we return to the implications of this assumption in the Discussion.

2.4. Complex enforcer

Finally, we can define a complex enforcer E_1 that modifies scenes by thinking about a mentalistic decider D_1. This model is identical to the simple enforcer (Section 2.2), with the only difference that it predicts the decider's behavior using the mentalistic decider model (Section 2.3), rather than the non-mentalistic one. That is, the term $p_{D_0}(a|s)$ from Eq. (1) is now replaced by $p_{D_1}(a|s, \phi)$ (i.e., the softmax of the utility function in Eq. (4)). This enforcer can therefore manipulate scenes under the assumption that the decider will attempt to decode their preferences.

2.5. Model implementation details

The computational framework specified above captures the proposal that people extract costs from physical scenes (i.e., what is the cost of taking different actions, and what costs did another agent incur in positioning the objects), and use them to make mental-state inferences. Because past work has already studied how people might infer costs from physical scenes (Yildirim et al., 2019), our interest is in testing how cost manipulations shape communication with objects. Therefore, in our model, we directly provide the costs associated with each scene change, which enables us to focus on the contribution of cost-based reasoning.

In principle, the parameters in our model can all be real-valued. For simplicity, we bounded costs and rewards to integers in the range 0 to 9. This enables us to easily interpret the range with 0 being null costs and rewards, and 9 being the highest possible costs or rewards that agents can have. We next set the cooperation parameter ϕ to take on integers between −25 and 25, which allows the model to consider extreme cooperative and adversarial cases (see Oey et al., 2022, for related work on adversarial mental-state reasoning). Our model code is available online at https://osf.io/57n4g.

3. Model analysis

Our first goal is to use our computational model to test whether our proposal can capture the emergence and use of low-cost communicative blockers (LCCBs). If our model failed to replicate this phenomena, this would imply that our account is incorrect. Specifically, our analyses consist of a set of simulations that test whether the enforcer and decider in our model can reproduce our target phenomena—creating and understanding LCCBs (inspired by those in Fig. 1).

To explore our model dynamics, we focused on the same simple domain with two doors—door A and door B—and an enforcer that

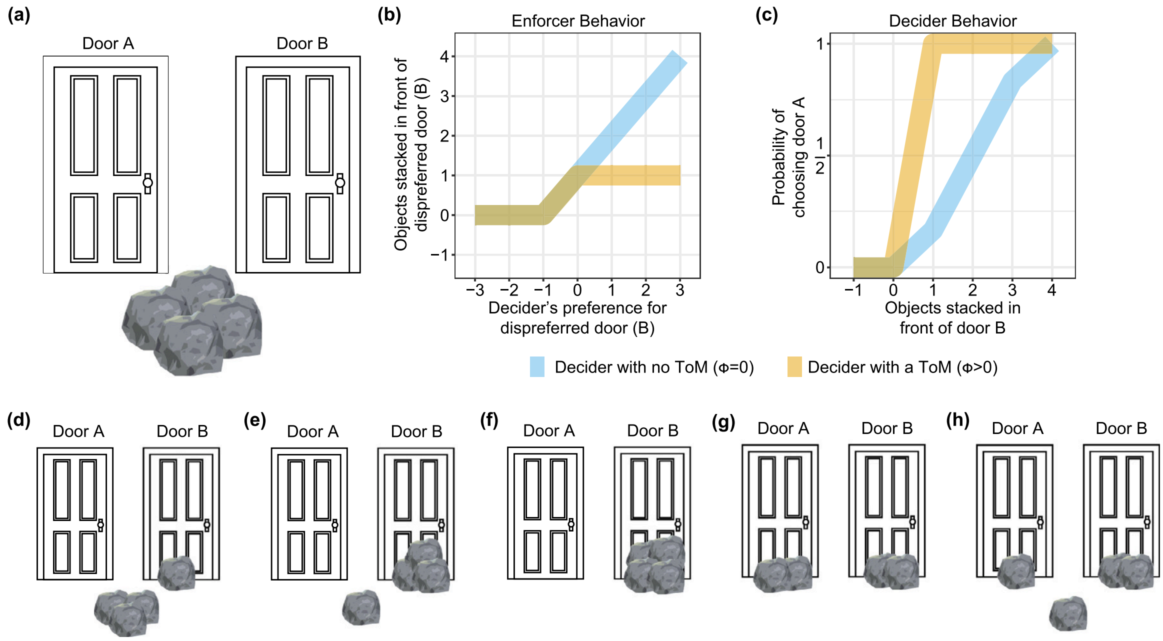

Fig. 2. Example event used to illustrate model performance. The environment consists of two doors and a stack of boulders between them. The enforcer's goal is to reposition the boulders to get the decider to choose door *A*. (a) Initial scene state. (b) Enforcer behavior. The *x*-axis shows the decider's preference for door *B* (negative values indicating a preference for door *A*) and the *y*-axis shows the number of boulders the enforcer stacks in front of door *B* (negative values indicating stacking objects on door *A*). The simple enforcer (blue line) builds the smallest possible physical barrier that will dissuade the decider. The complex enforcer (yellow line) places a single boulder in front of door *B*, even when the decider has a strong preference for going through it. (c) Deciders with a preference for door *B* reacting to boulders placed in front of that door. The non-mentalistic decider (blue line) slowly becomes more likely to choose door *A* as a function of how many boulders are blocking door *B*. The mentalistic decider (yellow line) recognizes the meaning of a single boulder and adjusts their behavior, immediately forgoing their preferred door *B* and choosing door *A* instead. (d–h) Visualization of some of the different scenes the enforcer could produce. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

wants the decider to choose door *A* (Fig. 2). We assume that the initial scene s_0 has a set of objects between the two doors, such that the objects do not initially block either door, and placing an object in front of either door is equally costly (Fig. 2a). For simplicity, we also assume that the cost the enforcer incurs in placing an object in front of a door is the same as the cost that the decider incurs when moving that object out of the way. To analyze the core dynamics of the model, we simulated a situation where the enforcer was maximally motivated to affect the decider's behavior (setting their reward to 9), and where the decider was also very cooperative (setting $\phi = 10$). We further set the softmax parameter to a minimum in order to remove any noise in the inferences (setting $\tau = 0.1$). We then tested our model's performance by varying the decider's relative preference for different options.

Fig. 2a shows a visual depiction of the initial state, and Figs. 2d–h shows five possible scene transformations that the enforcer could produce (stacking one, three, or four objects in front of door *B* alone, or also stacking any number of objects in front of door *A*). To understand our model behavior, we began by contrasting the simple and complex enforcers. Fig. 2b shows how many objects each enforcer chooses to stack in front of door *B* (the door they hope the decider will avoid) as a function of the decider's preference for this door. When the decider already prefers door *A* (negative decider preference along the *x*-axis in Fig. 2b), neither enforcer moves any objects. This reflects the enforcers' confidence that the decider will take door *A*, making any involvement unnecessary. When the decider prefers door *B* (positive decider preference along the *x*-axis in Fig. 2b), the enforcers begin to place objects in front of the door, producing two different types of behavior. The simple enforcer expects the decider to choose a door based only on their egocentric costs (how difficult is it to walk through each door?) and rewards (how much does the decider want to walk through each door?). Consequently, this enforcer stacks the minimum number of objects necessary to push the decider's choice towards door *A*. This is captured in Fig. 2b, where the blue line shows how the simple enforcer stacks more objects as the decider's preference becomes stronger. This behavior reflects a non-communicative barrier-building strategy, where the enforcer is attempting to make it just hard enough for the decider to cross through door *B*, with the hope that this added cost will shift their preference towards door *A*.

In contrast to the barrier-building strategy from the simple enforcer, the complex enforcer places a single object in front of door *B* (as in Fig. 2d), even when the decider really prefers that door (yellow line in Fig. 2b). We interpret this as the kind of communicative strategy that we aim to explain (reminiscent of Fig. 1): The strategy succeeds not because it imposes a high cost on deciders, but because it efficiently reveals the enforcer's mental states. In these cases, the enforcer knows that the decider's egocentric utilities will favor door *B*, because the single object imposes a negligible cost. The enforcer nonetheless chooses to place a single object in front of door *B* because they believe that the decider will infer that they are supposed to take door *A* instead.

Returning to our motivating examples (Fig. 1), this behavior resembles actions like placing a plastic chair to mark that a parking spot is taken. Here, a plastic chair does little to prevent someone from using the parking spot: moving the chair out of the way is easy, and the cost is probably insufficient to overcome a driver's desire to find a parking spot. However, the object is effective because it reveals that whoever placed the chair is requesting that their parking spot be respected.

Fig. 2c shows the behavior of our decider model. The non-mentalistic decider responds to the physical costs alone, becoming more likely to abandon their preferred door as a function of how many objects are blocking it. This is visualized by the blue line in Fig. 2c, which shows a continuous preference change as a function of the number of objects blocking their preferred door. By contrast, the mentalistic decider shows a sharp discontinuity: A single object in front of their preferred door is enough for them to understand that they should avoid that door. This is visualized by the yellow line in Fig. 2c, where the decider shows a rapid change in strategy as soon as a single object is in front of their preferred door. Together, these results show how our model gives rise to enforcers who use objects in a communicative manner and deciders who can infer the communicative meaning of these objects.

4. Experiment 1: Quantitative model evaluation

Having established that our account can replicate the qualitative use and recognition of low-cost communicative blockers (LCCBs), we next test whether our model's exact inferences match human intuitions. That is, our model predicts quantitative patterns about how strong people's intuitions should be in different displays. If participants can interpret LCCBs, but do so in a different way than our model does, the resulting large discrepancies between our model inferences and participant judgments would falsify our account.

In Experiment 1, participants saw a two-dimensional gridworld of a fruit farm with an entrance, pomegranate groves, pear groves, and a set of boulders placed by a farmer to protect their pomegranates from nearby hikers (farmers corresponding to enforcers and hikers to deciders from our Computational Framework).

We tested participants in two conditions (Fig. 3). In the *non-mentalistic* condition, hikers believe that the boulders are natural constraints, devoid of social meaning, and farmers plan for how many boulders to place accordingly. We therefore expect participants to infer that, the more boulders the farmer places, the more she believes that hikers want to take the pomegranates. We model this condition using the non-mentalistic decider model (Section 2.1) and the simple enforcer model (Section 2.2).

In our second condition, the *mentalistic* condition, hikers will always know that the boulders were placed by a farmer, and use the costs imposed by these objects to infer the farmer's preferences. In this condition, a single boulder does not necessarily imply an expectation that hikers do not like pomegranates (as would be implied in the *non-mentalistic* condition). Instead, a single boulder might reveal that the farmer expects hikers to infer that they should stay away and act accordingly. By contrast, if the farmer placed multiple boulders, this would reveal that she expects hikers to prefer pomegranates and be uncooperative (otherwise, a single communicative boulder would have sufficed). We model this condition using the mentalistic decider model (Section 2.3) and the complex enforcer model (Section 2.4).

All studies were approved by Yale's IRB (protocols "Culture and Cognition" #2000022403 and "Online reasoning" #2000020357). Data collection was obtained in the following experiment order: 3a (meaning inference), 3b, 2, 4, 1, 1 replication, 3a (unusualness ratings), and 3c. Our experimental procedure, stimuli, data, analyses, and pre-registrations (for Experiments 1 replication, 2, 3a, and 3c) are available at https://osf.io/57n4g. This manuscript includes all experiments, manipulations, and measures in this line of research.

4.1. Participants

80 U.S. participants (as determined by their IP address) were recruited using Amazon Mechanical Turk (n = 40 per condition; M = 34.81 years, SD = 10.31 years).

4.2. Stimuli

Stimuli consisted of 27 10-by-10 gridworlds, with two fruit groves (pears and pomegranates), a hiker, and a set of boulders (see Figs. 5a–c for examples). The stimuli were designed by parametrically varying two factors: the distance between the hiker and the groves (i.e., the natural cost of the environment; 5, 7, and 9 squares away) and the number of boulders blocking the pomegranates (i.e., the artificial cost introduced by the farmer; 1, 2, or 3 boulders). The hiker's starting position was randomly selected to be at one of the four corners, and the fruit groves were randomly placed on the two adjacent corners relative to the hiker.

4.3. Procedure

Participants read a brief cover story explaining that they would see hikers in different farmlands with pear and pomegranate groves (see Fig. 3 for paradigm schematic). The farmers, who were absent, did not mind hikers taking pears but they wanted to protect their pomegranates. To achieve this, farmers placed boulders in front of their pomegranate groves (see Figs. 5a–c for examples). Participants then completed a multiple-choice five-question quiz (see online OSF repository for questions) to ensure they understood the task. Participants that answered at least one question wrong were sent to the beginning of the cover story to try again. Participants that failed the questionnaire twice were not permitted to participate in the study.

Participants in the *non-mentalistic* condition were told that hikers thought the boulders were natural constraints, and that farmers planned how many boulders to place accordingly. That is, the farmer expected hikers to realize that the boulders make it harder to reach a fruit grove, but assume that this was simply a feature of the terrain, rather than an intentional design. In each trial, participants saw an arrangement of boulders and they were asked how much the farmer expected hikers to like pomegranates ("How much does this farmer think that hikers like pomegranates?", using continuous sliders ranging from "not at all" to "very much").

Participants in the *mentalistic* condition were told that hikers would always know that a farmer placed the boulders intentionally, and that farmers planned how many boulders to place accordingly. That is, the farmer expected hikers to know that the boulders make it harder to reach a fruit grove, and that these boulders were placed intentionally by someone. In each trial, participants saw an arrangement of boulders and they were asked how much the farmer expected the hiker to like pomegranates. In addition, because the complex enforcer and mentalistic decider include a cooperation parameter ϕ (i.e., the adopted utility weight; Powell, 2022), participants were also asked whether the farmer expected the hiker to be cooperative ("How cooperative does this farmer think hikers are?", using a continuous slider ranging from "not at all" to "very much").

All participants completed the same 27 trials (trial order randomized across participants), where we varied both the initial cost of obtaining each type of fruit (by manipulating the initial distance from the hiker) and the number of boulders that the farmer added (ranging from 1 to 3; see Stimuli).

4.4. Model predictions

Our model's parameters were set prior to data collection (and reflected in the pre-registration of the Experiment 1 replication; see Section 2.5 and SM for details). For each dependent variable in our task we computed our model's posterior predictive distribution, and used the expected value as the final model prediction.

4.5. Results

Our model and participant judgments showed an overall correlation of $r = 0.97$ ($CI_{95\%}$: 0.95–0.98; Fig. 4a). A pre-registered replication of this study produced identical results ($r = 0.98$; $CI_{95\%}$: 0.96–0.98; see SM for details). The fact that our model captures the fine-grained structure of people's inferences suggests that their inferences resembled the ones obtained by reasoning about the farmer's desires via recursive social reasoning.

Fig. 5 shows three example trials that highlight the inferences that our model and participants made. In Figs. 5a–c, the hiker's distance to pomegranates and pears is matched (making the initial cost identical) and the number of boulders in front of the pomegranates varies from 1 to 3. In the *non-mentalistic* condition, the number of boulders should reveal how much the farmer thinks hikers will want to get the pomegranates (because the purpose of the boulders is to introduce a

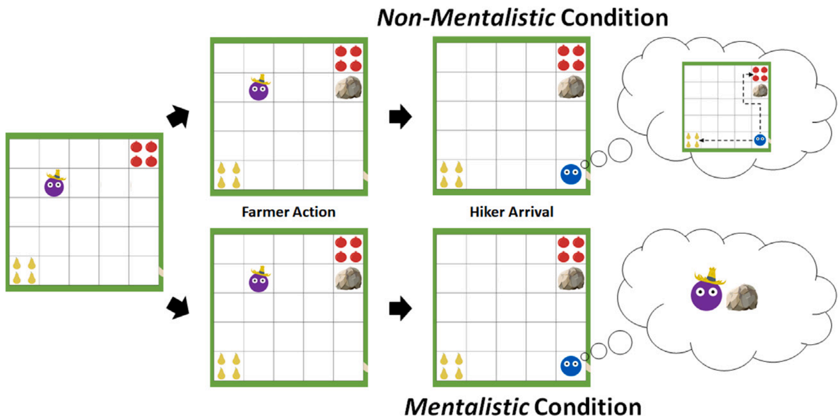

Fig. 3. Visual schematic of Experiment 1 cover story. Participants learned that a farmer (purple agent) wanted to protect their pomegranates and placed boulders to block the way before leaving. After leaving, a hiker would arrive and decide which fruit to take. In the *non-mentalistic* condition, the hikers treat the boulders as natural constraints, and they therefore decide what to do without thinking about the farmer. In the *mentalistic* condition, the hikers know that a farmer must have placed the boulders, and use this to infer what to do. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

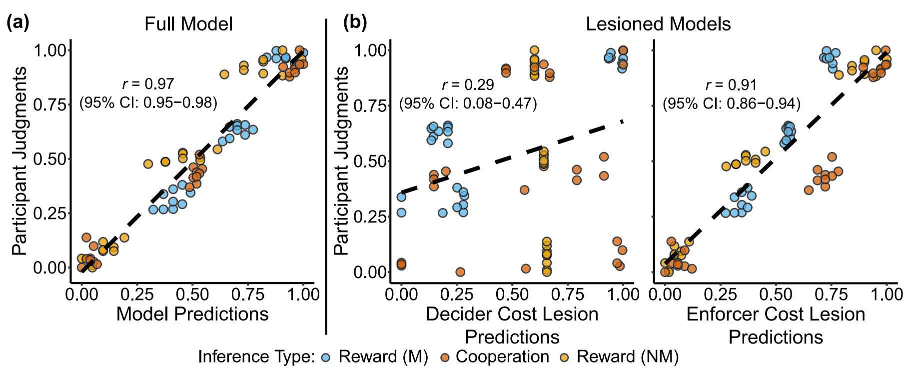

Fig. 4. Experiment 1 results. Each point represents a judgment, with model predictions on the *x*-axis and participant judgments on the *y*-axis. Participants in the *non-mentalistic* condition (NM) condition inferred the hiker's expected preferences, and participants in the *mentalistic* condition (M) additionally inferred the hiker's expected cooperativeness (ϕ). (a) Correlation between our full model and participant judgments. (b) Correlation between lesioned models and participant judgments. Model lesions include removing the influence of cost from the decider (left) and removing the influence of cost from the enforcer (right).

physical cost that outweighs the hikers' desires). Figs. 5d–f show this effect in both our model and participant reward inferences (with each plot corresponding to the stimuli directly above it; e.g., panel (d) corresponding to map (a)). As the number of boulders increased, participants and our model inferred a stronger preference for pomegranates.

In the *mentalistic* condition, the boulders not only impose a physical cost, but allow hikers to infer the farmer's preferences. Therefore, a single boulder (Fig. 5g) does not necessarily imply that hikers must not like pomegranates that much (as it did in the *non-mentalistic* condition; Fig. 5d). Instead, the farmer may have used a single boulder to reveal that they did not want hikers to take the pomegranates. Consistent with this, both participants and our model inferred that hikers could have a higher desire for pomegranates (compare Figs. 5g and 5d), but were highly cooperative. That is, participants and our model inferred that a single boulder was effective because it revealed the farmer's preferences to cooperative hikers (despite its cost not being high enough to outweigh their preferences). This reward difference across conditions is further visualized in Fig. 6, and was significantly different across conditions ($\Delta R = 0.24$; $p < 0.001$ from a two-tailed t-test).

When the number of boulders blocking the pomegranates increases (Figs. 5h–i), the farmer's additional actions (placing more than one boulder) can be explained by inferring that hikers must really want the pomegranates and not be particularly cooperative (given that they will be able to infer that the farmer wants them to stay away). Consistent with this, both our model and participants infer a stronger hiker desire and a lower cooperativeness as the number of boulders increases (see SM for additional results of a linear mixed-effects regression predicting these participant reward inferences as a function of boulder count and condition).

4.5.1. Alternative models

While our model captured participant inferences with quantitative accuracy, it is possible that participants reached similar inferences through simpler mechanisms. To test this, we considered two alternative models. A first possibility is that people focus only on an object's position, without considering the costs that it might impose on observers. In our experiment, this means that hikers do not consider the cost of navigating around boulders. We call this model the *Decider Cost Lesion* as it is similar to our model with the difference that it does not reason about the cost that objects impose on deciders. Fig. 4b (left) shows how this model was no longer able to explain participant judgments ($r = 0.29$; $CI_{95\%}$: 0.08–0.47), and was also reliably worse than our main model ($\Delta r = 0.68$; $CI_{95\%}$: 0.49–0.89). This result confirms that the cost imposed on deciders is critical for capturing human-like inferences.

A second possibility is that people do consider the costs that an object imposes on their actions (e.g., detecting that an object is making it harder for them to get a certain fruit), but they do not consider the effort that someone had to incur in positioning the object. In our experiment, this means that people do not think about the cost farmers incur when placing boulders. We call this model the *Enforcer Cost Lesion* as it is similar to our model with the difference that it does not reason about the cost the enforcer incurs. Fig. 4b (right) shows how this lesioned model compares to participant judgments. Although

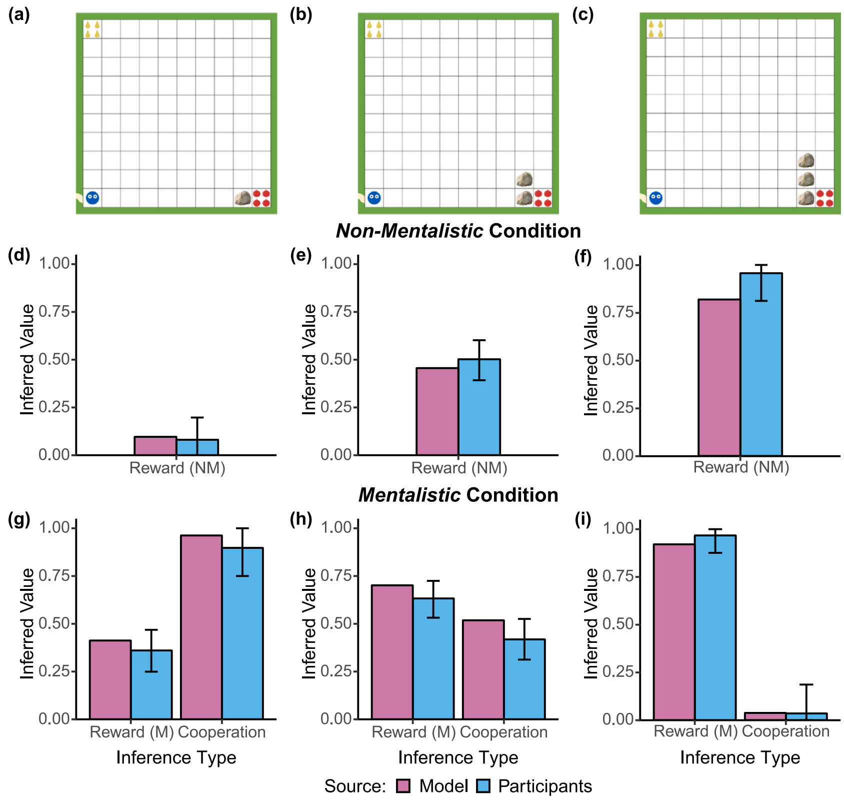

Fig. 5. (a–c) Example stimuli from Experiment 1. In these examples, both fruit groves were equally far and only varied on the number of boulders a farmer placed. (d–f) Model predictions and participant judgments from the *non-mentalistic* condition in purple and blue, respectively. (g–i) Model predictions and participant judgments from the *mentalistic* condition in purple and blue, respectively. Inference type is along the *x*-axis and the inferred value is along the *y*-axis. Error bars are bootstrapped 95% CIs. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

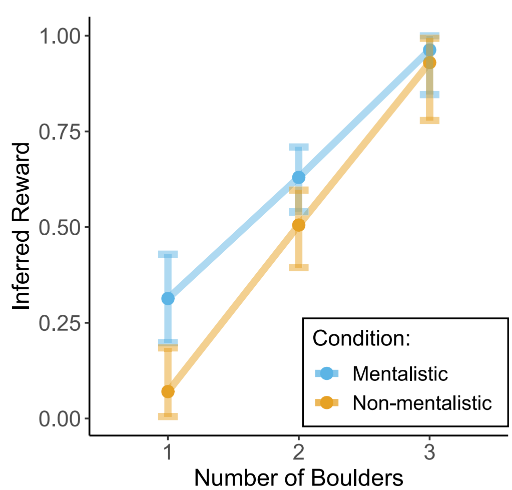

Fig. 6. Reward inferences across the *non-mentalistic* condition and *mentalistic* condition in Experiment 1. The number of boulders placed by the farmer is on the *x*-axis and the reward participants inferred is on the *y*-axis. Error bars are bootstrapped 95% CIs.

this model performed worse than our main model ($\Delta r = 0.06$; $CI_{95\%}$: 0.03–0.10), it was nonetheless able to capture the pattern of inferences about deciders quite well ($r = 0.91$; $CI_{95\%}$: 0.86–0.94).

These results suggest that participant inferences may not depend as heavily on the cost incurred by the enforcer (i.e., the farmer). We believe this is intuitive for the situations that we focused on. For instance, when encountering a broom positioned directly across a door, we intuitively focus on the cost that the broom imposes on us, rather than thinking about the cost the enforcer incurred. At the same time, Fig. 4b (right) reveals that this lesioned model nonetheless fails to capture a subset of participant intuitions that our main model was able to capture. Specifically, this lesioned model over-estimated hikers' cooperativeness when compared to humans in trials with two boulders (visualized as a cluster of orange points that falls most distant from the best-fit line in Fig. 4b, right). This is because, according to this lesioned model, placing three boulders is as easy as placing two boulders. Therefore, the farmer choosing not to place an extra boulder at no cost would only be reasonable if the hikers were highly cooperative, to the point that placing a single boulder was guaranteed to be as effective as placing more boulders to block the way. Together, this analysis suggests that the cost incurred by enforcers is less critical for capturing how we infer the communicative meaning of an object, but that people are nonetheless sensitive to it, and use it to infer other agents' cooperativeness. Overall, these lesions show how considering the cost that enforcers incur in positioning objects, and how these objects also impose a cost on deciders, are key to explaining how participants reasoned about objects in our experiment.

5. Overview of experiments 2–4

Our model analyses and Experiment 1 show that people can derive inferences that are quantitatively similar to those from our model. While these results show that people can make these types of social inferences, they do not imply that this is what people do when encountering communicative objects. In Experiments 2–4, our goal is therefore to test an alternative hypothesis: Could simple conventions without inference explain the use of communicative objects in their entirety? That is, conventional knowledge is undoubtedly critical to the everyday use of communicative objects. Our goal is therefore not to question its importance, but to ask what happens when conventional knowledge is unavailable, such as when we encounter an unfamiliar object that

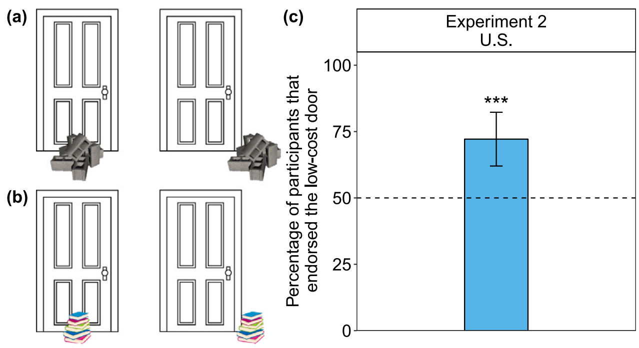

Fig. 7. (a–b) Example stimuli from Experiment 2. In both panels, the left side shows a low-cost door and right side shows a no-cost door. (c) Experiment 2 results. The bar represents the percentage of participants that associated the low-cost door with having a communicative purpose. Error bars are bootstrapped 95% CIs.

might have a communicative purpose. In these cases, do people rely on social inferences like the ones we proposed? Or are they unable to make any conclusions given the absence of an explicit convention?

Given that people will use conventional knowledge when it is available, our experiments here focus on objects that are not associated with a pre-existing communicative meaning. We begin by testing two predictions. First, if the meaning of communicative objects were based on explicit convention alone, then people should detect an object as communicative only when they have been explicitly taught about its meaning (therefore falsifying our account). By contrast, our account predicts that people should be more likely to associate low-cost novel objects with a communicative purpose, relative to novel objects that impose no cost (as these fail to reveal the mental states of whoever positioned the object). We test this prediction in Experiment 2. Second, if the meaning of objects were driven by explicit pedagogy and convention alone, then people should be unable to infer the meaning of a novel object, even when they know that the object has a communicative purpose (therefore falsifying our account). By contrast, our account predicts that people should be able to infer the communicative meaning of an object when its placement (i.e., the cost it imposes) reveals the enforcer's mental states. We test this prediction in Experiment 3. Finally, if people are engaging in social inferences to infer the meaning of novel objects, our work brings forth the question of how quickly these meanings might become conventionalized. In Experiment 4, we test the idea that people might be able to quickly treat the meaning of a novel object as conventional.

6. Experiment 2: Are objects that impose a cost more likely to be communicative?

If people's reasoning about low-cost communicative blockers is driven entirely by explicit object-meaning conventions, then people should report that an object is communicative only when they have been explicitly taught its meaning. In Experiment 2, we therefore tested whether people believe that objects that impose a low cost on deciders are more likely to be communicative relative to objects that do not impose a cost, as our account predicts but the explicit convention account does not.

6.1. Participants

80 U.S. participants (as determined by their IP address) were recruited using Amazon Mechanical Turk (M = 37.84 years, SD = 12.22 years). 14 additional participants were recruited and replaced for failing our inclusion criteria (see Results).

6.2. Stimuli

Stimuli consisted of eight images of pairs of doors, with each pair consisting of a "low-cost door" and a "no-cost door" (e.g., Figs. 7a–b). Each of the eight pairs was associated with one of eight objects that are not conventionally used to communicate: a plant, a chair, a pile of books, a pile of cinderblocks, some tape, some meter sticks, a hat, and a fishbowl tied to a tack on a door frame (see online OSF repository for the full stimuli set). In the low-cost doors, the object was placed directly in the middle of the doorway (e.g., Figs. 7a–b, left), and in the no-cost doors, the object was placed next to the door, not blocking the way (e.g., Figs. 7a–b, right). Half of the door pairs were open and the other half were closed.

6.3. Procedure

Participants were asked to imagine leaving an office and encountering a pair of doors, each with an object nearby. Participants then answered a simple multiple-choice attention-check question ("What objects are in front of the doors?"). Participants that answered incorrectly were sent to the beginning of the cover story and not permitted to access the experiment until they answered correctly.

Participants then saw a single trial containing a low-cost door and a no-cost door, both with the same object nearby (see Figs. 7a–b for examples; door order randomized across participants). Participants were asked, "Which door do you think someone was trying to tell you something?", followed by a manipulation check ("Which door requires more work to walk through?") and an inclusion question ("Do you think you would be able to walk through this door if you wanted to?"). These questions were always presented in the same order (see online OSF repository for the full procedure details).

6.4. Results

Participants who did not think they could walk through the doors were excluded from the study and replaced (as our interest is in the inferences people make when objects are not seen as insurmountable physical constraints; n = 14; 14.89% exclusion rate). Of our final sample, 72.50% of participants reported that the low-cost door was more likely to be communicative ($CI_{95\%}$: 62.50%–82.50%; $p < 0.001$ from a two-tailed binomial test; Fig. 7c), rather than performing at chance, as expected by the explicit pedagogy account (see SM for a supplemental analysis confirming this result). Our exclusion rate (14.89%) was lower than recent estimates of attentiveness on Mechanical Turk (estimated to be at approximately 20%; Arechar and Rand, 2021), suggesting that these participants were simply inattentive. However, these participants showed the same qualitative pattern of responses as those included in the task (see SM for details on excluded participants). It is therefore possible that these participants were attentive but did not interpret our inclusion question ("Do you think you would be able to walk through this door if you wanted to?") as referring to physical plausibility alone, integrating social expectations as well (given the heavy social focus of the task).

7. Experiment 3a: Are low-cost objects interpreted as communicative blockers?

Having found that people are more likely to interpret a low-cost object as communicative, in Experiment 3a, we next test what meaning people are more likely to associate with it. If the object is unfamiliar, the explicit convention account predicts that participants should be at a loss about what it means, given the absence of an established object-meaning mapping. Instead, if people are making inferences about why someone would place the object to impose a cost, they should infer that the object is more likely to mean that they should avoid the door.

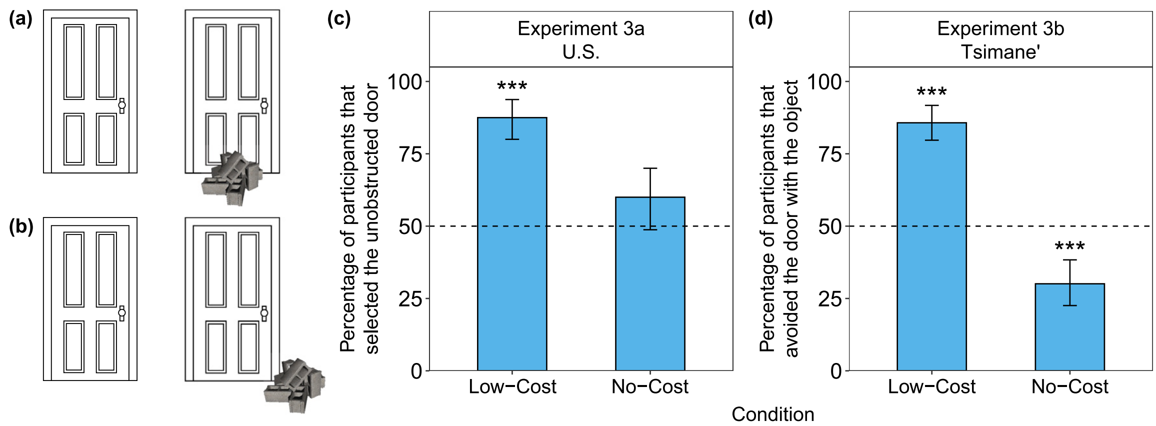

Fig. 8. (a) Example stimuli from the *low-cost* condition in Experiment 3a. (b) Example stimuli from the *no-cost* condition in Experiment 3a. (c) Experiment 3a results. The bars represent the percentage of U.S. participants that selected the empty door as a function of condition. (d) Experiment 3b results. The blue bars represent the percentage of Tsimane' participants that inferred that they should not go through the door as a function of door type. Error bars are bootstrapped 95% CIs.

7.1. Participants

160 U.S. participants (as determined by their IP address) were recruited via Amazon Mechanical Turk ($n = 80$ per condition; $M = 34.85$ years, $SD = 8.38$ years). 17 additional participants were recruited and replaced for failing our inclusion criteria (see Results).

7.2. Stimuli

Stimuli consisted of 16 images of pairs of doors, with each pair consisting of an empty door and a door with an object nearby (e.g., Figs. 8a–b; using the same objects from Experiment 2). In the "low-cost pair", the object was placed directly in the middle of one of the doorways (e.g., Fig. 8a), and in the "no-cost pair", the object was placed next to one of the doors, not blocking the way (e.g., Fig. 8b). Half of these door pairs were open and the other half were closed.

7.3. Procedure

The procedure was similar to Experiment 2. Participants were asked to imagine leaving an office building and finding two identical exits. Participants then saw that one of the doors had an object nearby, and they were told that it was unclear whether someone wanted them to take that door or to avoid it. Participants were then asked a multiple-choice attention-check question: "What is the only difference between the two exits?" Participants next saw a pair of doors (either a low-cost pair for participants randomly assigned to the *low-cost* condition or a no-cost pair for participants randomly assigned to the *no-cost* condition; door order randomized across participants), and were asked: "What do you think someone was trying to tell you about the door with the OBJECT?" (possible responses: "You should walk through the door with the OBJECT" or "You should not walk through the door with the OBJECT"). Participants then responded to the same manipulation-check and inclusion questions from Experiment 2.

7.4. Results

Participants who said the empty door was harder to walk through were excluded from the study and replaced ($n = 17$; 9.60% exclusion rate). Of our final sample, 87.50% of participants in the *low-cost* condition inferred that they were supposed to avoid the door with the object ($CI_{95\%}$: 80.00%–93.75%; $p < 0.001$ from a two-tailed binomial test). By contrast, only 60.00% of participants in the *no-cost* condition inferred that this door should be avoided ($CI_{95\%}$: 48.75%–70.00%; $p = 0.093$ from a two-tailed binomial test[1]), a proportion not significantly different from chance. Moreover, the number of participants inferring that they should avoid the door was significantly higher in the *low-cost* condition relative to the *no-cost* condition ($p < 0.001$ by Fisher's exact test). The fact that participants did not perform at chance in both conditions suggests that people do not rely purely on conventional knowledge. The pattern of results from our excluded participants was qualitatively consistent with that of participants included in the task (see SM for details on excluded participants).

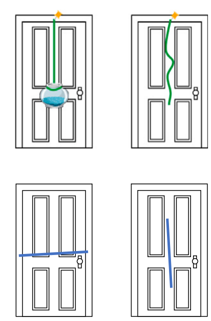

Fig. 9. Two pairs of doors from Experiment 3a, where both the *low-cost* (first column) and *no-cost* (second column) versions had the object in front of the door. Consistent with the main results, participants significantly inferred avoidance in the *low-cost* versions but not the *no-cost* versions. This suggests that having an object in front of a door is insufficient for triggering these inferences.

Our experiments so far are consistent with our proposal: participants may be engaging in cost-based reasoning when interpreting the meaning of low-cost communicative blockers. However, there are two alternative explanations that are important to consider. First, is it possible that the position of an object broadly indicates what space it concerns, independent of the cost it imposes? That is, rather than participants reasoning about object costs, they may arrive at similar inferences by seeing any inapposite object in front of a door.

To test this possibility, we conducted a secondary analysis where we separated trial types within our experiment. Our overall experimental design used eight objects in a *low-cost* and a *no-cost* condition. For two of the eight objects, both the *low-cost* and *no-cost* variants had the object in front of the door (Fig. 9). If the inapposite account is correct,

[1] Our pre-registered analysis proposed to use logistic regressions to study this effect. We instead present binomial tests for clarity, but the results are identical under our pre-registered analysis and can be found in SM.

then participants should infer avoidance in both the *low-cost* and *no-cost* versions for these two objects. In these trials, participants inferred avoidance in the *low-cost* condition significantly above chance (19/20; $p < 0.001$ from a two-tailed binomial test), but were at chance in the *no-cost* condition (14/20; $p = 0.120$ from a two-tailed binomial test). This suggests that having an object in front of a space is insufficient for triggering these inferences.

A second possibility is that people simply believe that more 'unusual' arrangements of objects are more likely to signal avoidance. To answer this question, we conducted an additional experiment where we asked participants to rate how unusual each arrangement of objects appeared to them, and we repeated our analyses while controlling for unusualness. In this analysis, we found that the cost that an object imposes continued to significantly predict participant answers, independent of the object's unusualness (see Section 5 in SM).

8. Experiment 3b: Replication with the Tsimane'

Experiment 3a suggests that people's reasoning about low-cost communicative blockers cannot be reduced to explicit object-meaning conventions. However, it is possible that these inferences are culture-specific. Because our model is built on simple aspects of human cognition that are thought to be universal, the absence of these inferences in other cultures would challenge our account. As a first step in exploring this possibility, we replicated a variation of Experiment 3a with the Tsimane'—a farming-foraging group native to the Bolivian Amazon. The Tsimane' live in non-industrialized communities along the Maniqui river and have less exposure to market-integrated communities compared to U.S. participants. Comparing the Tsimane' and WEIRD participants (Western, educated, and from industrialized, rich, and democratic countries; Henrich et al., 2010) has helped identify cultural influences in color-word vocabulary (Conway et al., 2020; Gibson et al., 2017) and music perception (McDermott et al., 2016), and has also helped rule out cultural influences in other domains, such as the stages of number-word learning in children (Jara-Ettinger et al., 2017; Piantadosi et al., 2014) and the ways in which people identify communicative action (Royka et al., 2022). We therefore sought to test the Tsimane' as a way to explore if these inferences also emerge in a culture that is substantially different from the U.S.

8.1. Participants

133 Tsimane' adults were recruited in their local communities in the Bolivian Amazon ($M = 33.12$ years, $SD = 15.40$ years). 17 additional participants were recruited but excluded from the study for failing to complete the study (see Results).

8.2. Stimuli

Stimuli consisted of six images of doors, each with an object in front of it (e.g., Figs. 7a–b). We used a subset of the objects used in Experiment 3a that Tsimane' participants were familiar with (as determined by our interpreters) while remaining unconventional as communicative objects: a plant, a chair, and a pile of cinderblocks. Each object was associated with two different doors: a "low-cost door", where the object was placed directly in the middle of the doorway (e.g., Fig. 7a, left), and a "no-cost door", where the object was placed next to the door, not blocking the way (e.g., Fig. 7a, right). Half of these doors were open and the other half were closed.

8.3. Procedure

The procedure was adapted from Experiment 3a to be more intuitive for our participants, based on feedback from our interpreters. Participants were asked to imagine deciding to enter a friend's house through one of two possible doors. Participants were then shown a low-cost door and a no-cost door sequentially (order counterbalanced across participants, with different objects used for each door) and were asked: "Do you think the owner wants you to enter or stay away?"

8.4. Results

Participants that failed to complete both trials were excluded from the study ($n = 17$; 11.33% exclusion rate; see SM for details on excluded participants). Like U.S. participants, Tsimane' participants inferred that they should avoid a door when the object was minimally blocking the door (85.71%; $CI_{95\%}$: 79.70%–91.73%; $p < 0.001$ from a two-tailed binomial test), but not when the object was on the side of the door (30.08%; $CI_{95\%}$: 22.56%–38.35%; $p < 0.001$ from a two-tailed binomial test).

8.5. Discussion

These results suggest that, like U.S. participants, Tsimane' participants also inferred avoidance from objects that impose a low cost. Critically, this experiment used objects that were familiar to the Tsimane', but not typically used by them to communicate (as determined by our local interpreters). This approach followed the same logic as our design with U.S. participants, which also used familiar objects but, critically, not ones typically used to communicate. This enabled us to maximize *equivalence* (Matsumoto & Yoo, 2006; Poortinga, 1989; Van de Vijver & Leung, 2021)—the goal of reaching similarity in conceptual meaning across groups to support meaningful comparisons. At the same time, a stronger test of our hypothesis would have included entirely novel objects, which would have allowed us to test the nature of these inferences without any possible influence from prior object knowledge. These results, therefore, only provide evidence that people can infer the communicative meaning of familiar objects that are not typically communicative, and we do not know if these inferences would extend to entirely novel objects.

9. Experiment 3c: Inferences from conventional communicative blockers

In Experiments 3a and 3b, we found evidence that people are sensitive to the cost an object imposes when reasoning about its potential communicative meaning. Importantly, these experiments used objects with no pre-existing communicative meaning associated with them. Under our account, these inferences become critical when people do not have a pre-existing convention, but may become less important when they already know an object's communicative meaning. In Experiment 3c, we replicated Experiment 3a using conventional objects. If people are constantly making cost-based inferences with all communicative objects, these results should replicate the pattern of Experiment 3a: inferring avoidance in the *low-cost* condition, but not in the *no-cost* condition. However, if these inferences are only at work when encountering novel objects, people should report the conventional communicative meaning of the object regardless of the cost that it imposes.

9.1. Participants

60 U.S. participants (as determined by their IP address) were recruited via Prolific ($n = 80$ per condition; $M = 35.72$ years, $SD = 12.48$ years). 4 additional participants were recruited and replaced for failing our inclusion criteria (see Results).

9.2. Stimuli

Stimuli consisted of six images of pairs of doors, with each pair consisting of an empty door and a door with an object nearby (similar to those used in Experiment 3a, but using different objects; e.g., Figs. 10a–b). Here we used objects that are conventionally used as communicative blockers: a traffic cone, construction tape, and a stanchion.

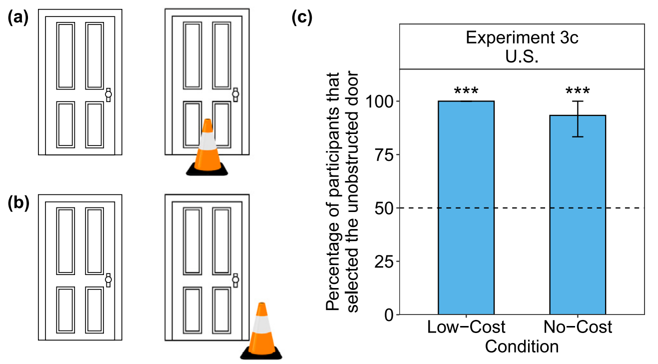

Fig. 10. (a) Example stimuli from the *low-cost* condition in Experiment 3c. (b) Example stimuli from the *no-cost* condition in Experiment 3c. (c) Experiment 3c results. The blue bars represent the percentage of U.S. participants that selected the empty door as a function of condition. Error bars are bootstrapped 95% CIs (not visible in *low-cost* condition due to participants performing at ceiling).

9.3. Procedure

The procedure was identical to that used in Experiment 3a, with the only difference being the objects that participants saw. Participants then responded to the same manipulation-check and inclusion questions from Experiment 3a.

9.4. Results

Participants who said the empty door was harder to walk through were excluded from the study and replaced (n = 4; 6.25% exclusion rate), since the empty door is never harder to walk through (see SM for details on excluded participants). Of our final sample, 100.00% of participants in the *low-cost* condition inferred that they were supposed to avoid the door with an object ($CI_{95\%}$: 100.00%–100.00%; $p < 0.001$ from a two-tailed binomial test.[2] 93.33% of participants in the *no-cost* condition also inferred that the door should be avoided ($CI_{95\%}$: 83.33%–100.00%; $p < 0.001$ from a two-tailed binomial test. While the number of participants inferring that they should avoid the door was qualitatively higher in the *low-cost* condition relative to the *no-cost* condition, this difference was not significant ($p = 0.492$ by Fisher's exact test). These results suggest that conventional knowledge was not driving our effect in Experiment 3a, and that people may have a lower reliance on costs when interpreting conventional communicative objects.

10. Experiment 4: Conventionalizing object meanings

Experiments 2, 3a, and 3b suggest that people are sensitive to costs when reasoning about an object with no pre-existing convention about its meaning. Experiment 3c further shows evidence that this sensitivity disappears when the object already has a conventional meaning associated with it. Experiment 4 presents an initial test of how objects that trigger inferences might become conventionalized (no longer requiring inference).

The process of associating objects with meaning might be particularly valuable, as it would help people minimize cognitive demands (Back & Apperly, 2010; Birch & Bloom, 2007; Keysar et al., 2000) and cognitive effort (Kool & Botvinick, 2018; Shenhav et al., 2017). That is, people might do social inference to interpret novel communicative objects, but treat them as conventional quickly afterwards. This hypothesis would help explain why some communicative objects, like chairs to indicate shoveled parking spots (e.g., Fig. 1c), are used consistently (although not always; e.g., Fig. 1i), and is parsimonious with a resource-rational view of the mind (Griffiths et al., 2015; Lieder & Griffiths, 2019).

Participants in Experiment 4 were first asked to infer the meaning of a low-cost object, and then completed a second trial that either showed a door with a picture of the same object (*congruent* condition) or a picture of a new object (*incongruent* condition). The critical idea in this experiment is that the picture in the second trial never imposes a cost. If participants interpret that picture independently, they should perform at chance when asked what it means (replicating the responses in the *no-cost* condition of Experiment 3a). Alternatively, people might infer the meaning of the low-cost object in the first trial and then immediately treat it as conventional. If so, then participants should report the same inference when they see a picture of the same object (i.e., in the *congruent* condition) despite the object not imposing a cost, but not when they see a picture of an unrelated object (i.e., in the *incongruent* condition).

10.1. Participants

160 U.S. participants (as determined by their IP address) were recruited using Amazon Mechanical Turk ($n = 80$ per condition; M = 36.13 years, SD = 10.90 years). 42 participants were recruited and replaced for failing our inclusion criteria (see Results).

10.2. Stimuli

Stimuli consisted of 16 images of pairs of doors, with each object (using the same eight objects from Experiments 2 and 3a) associated with two pairs. In the "low-cost pair", the object was placed directly in the middle of one of the doorways (e.g., Fig. 11a), next to an empty door. In the "symbol pair", a picture of the object was placed directly in the middle of one of the doorways (e.g., Figs. 11b–c), next to an empty door. Half of these door pairs were open and the other half were closed (with the picture hanging from the top of the door frame when the door was open).

10.3. Procedure

The experiment began in an identical way to Experiment 3a, with the difference that the cover story named the person who had positioned the communicative object. This was done so that we could ensure participants understood that the communicative objects came from the same agent across trials (name randomized across participants).

In the first trial, participants saw a low-cost pair of doors (e.g., Fig. 11a) and were asked: "What do you think NAME was trying to tell you about the door with the OBJECT?" (with the same possible responses from Experiment 3a) and "How confident are you that that's what NAME was trying to tell you?" (using a continuous slider ranging from "not confident at all" to "very confident").

Participants were told that their inference was correct (regardless of their answer) and they were next presented with a symbol pair of doors. In the *congruent* condition, one of the doors had a picture of the object from the previous trial (e.g., Fig. 11b after seeing Fig. 11a, right, on the first trial). In the *incongruent* condition, one of the doors had a picture of a new object (e.g., Fig. 11c after seeing Fig. 11a, right, on the first trial). Participants were asked the same two questions from the first trial, followed by the manipulation-check and inclusion questions from Experiments 2 and 3a.

[2] Our pre-registered analysis proposed to use t-tests to study this effect. We instead present binomial tests for clarity, but the results are identical under our pre-registered analysis and can be found in SM.

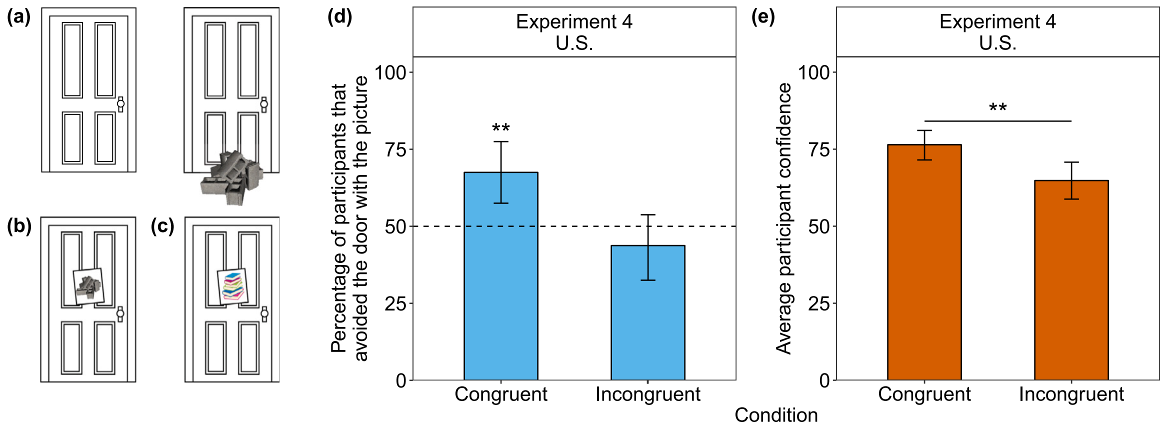

Fig. 11. (a) Example stimuli from the first trial of Experiment 4. (b–c) Example of a door from the *congruent* and *incongruent* condition relative to (a), respectively. (d) Percentage of participants that inferred that they should avoid the door with the picture as a function of condition in Experiment 4. (e) Average reported confidence rating in this inference. Error bars are 95% bootstrapped CIs.

10.4. Results

Participants who did not respond that the object in the first trial imposed a cost were excluded from the study and replaced ($n = 42$; 20.79% exclusion rate), as our interest is in how participants generalize the inferred meaning of objects that they perceived as imposing a cost (see SM for details on excluded participants).

The first trial replicated our results from Experiment 3a, with 66.25% ($CI_{95\%}$: 58.75%–73.75%) of participants inferring that they should avoid the door with the object ($p < 0.001$ from a two-tailed binomial test). Participants reported an average confidence rating of 76.96% ($CI_{95\%}$: 73.71%–80.03%).

We next turned to our main question of interest. If participants treat the picture in the second trial as a novel signal, they should perform at chance in both conditions, as the picture does not impose a cost. By contrast, if participants assume that the signaler was treating the object as a new convention, they should infer that the door should be avoided in the *congruent* condition, but perform at chance in the *incongruent* condition. As predicted, participants in the *congruent* condition judged that the door with the picture should be avoided, despite the picture not imposing a cost (67.50%; $CI_{95\%}$: 57.50%–77.50%; $p < 0.01$ from a two-tailed binomial test). In the *incongruent* condition, only 43.75% of participants inferred that the door should be avoided ($CI_{95\%}$: 32.50%–55.0%; $p = 0.314$ from a two-tailed binomial test). Moreover, participants were significantly more confident in their interpretation of the picture in the *congruent* condition (76.45%; $CI_{95\%}$: 71.46%–81.09%) relative to the *incongruent* condition (64.83%; $CI_{95\%}$: 58.78%–70.81%; $W = 3978$, $p < 0.01$ from a U-test).

11. Experiments 2–4 discussion

Experiments 2–4 suggest that people can infer the potential communicative meaning of an object in the absence of explicit pedagogy and convention. People's inferences were qualitatively consistent with our proposal, where communicative inferences are guided by reasoning about others' mental states. Is it possible that participants arrived to these inferences through simpler heuristics?

A first concern is that an object placed in front of a door might be more salient than an object placed to the side. It is possible that participants in Experiment 2 thought that visually-salient objects (rather than low-cost objects) were more likely to be communicative. However, that it is unclear why people would expect visual salience to imply avoidance, and there are cases where visual salience is interpreted to mean the opposite (Misyak et al., 2016). In addition, a general expectation that communicative objects should be visually salient is not mutually exclusive with our account, and is in fact consistent with the main idea in our proposal: To elicit mental-state inferences through objects, the communicator must also ensure that the recipient will notice the object in the first place.

A second potential concern is that people find unusual arrangements of objects to be more likely to be communicative, and to signal avoidance. To test for this possibility, in Experiment 3a we collected "unusualness" ratings for our stimuli and found that the cost information significantly explained participant judgments when controlling for unusualness (see SM for details).

12. General discussion

Human communication is remarkable in that, beyond words and gestures, we can communicate through objects. Here we proposed that this ability emerges from our capacity to represent and infer other people's mental states. Specifically, we proposed that, if we can reason about other people's mental states based on how they manipulate objects in the environment, then people can also arrange objects with the purpose of revealing their mental states to agents who encounter these objects.

In exploring this idea, our paper makes three contributions. First, we implemented a computational model of social inference from physical objects. This model revealed how communicators can use objects to elicit mental-state inferences in observers, and how observers can infer the communicative meaning of objects, without any direct communication occurring between the agents. This provided proof-of-concept that the computations proposed in our account are sufficient to give rise to this phenomena. Second, we directly tested whether people's inferences about the communicative meaning of an object could be explained by our account. Finally, we tested whether people's intuitions about communicative objects could only be the result of explicit systems of pedagogy and convention. Combined, these results suggest that mental-state inferences support the creation and interpretation of novel communicative objects, while pedagogy and convention drive their widespread use.

12.1. Model assumptions and study limitations

At its core, our model consisted of a Theory of Mind implementation which we built following several computational principles that each enjoy strong empirical support: social interactions involve agents thinking about each other's mental states (Frank & Goodman, 2012; Goodman & Frank, 2016); mental-state reasoning is structured around an assumption that agents maximize utilities (Jara-Ettinger et al., 2016; Jern et al., 2017; Lucas et al., 2014); inferences about other people's minds is performed via some approximation of Bayesian inference (Baker et al., 2017, 2009); and agents are typically cooperative, particularly when cooperation is easy (Powell, 2022; Rand, 2016; Warneken & Tomasello, 2006). To capture how we might extract social information from the physical world, we made three further assumptions in our experiments and model.

Our first assumption was that agents can identify which objects in a scene were manipulated by an agent (as opposed to being the result of some inanimate force; e.g., the wind). This assumption was reflected in the formulation of our decider model, where deciders know both the initial scene s_0 and the final scene s (such that any discrepancy between s and s_0 reveals the costs that the enforcer incurred, and the costs that they introduced to the decider). While related research shows that some form of this capacity emerges in infancy (Keil & Newman, 2015; Newman et al., 2010), it is likely that in more realistic contexts people do not know the initial scene s_0 and instead have a prior distribution over potential initial scenes $p(s_0)$.

Our second assumption was that the objects we considered were intentionally placed by an agent. In real-world situations, this is not known a priori and people must determine which objects were placed intentionally and which were not. Intuitively, there are many objects that, when encountered in front of a door, would not elicit a communicative inference because their placement appears unintentional. For instance, finding an empty paper bag (which an agent might have simply discarded), a soccer ball (which might have rolled over to the front of the door), or a wallet (which an agent might have dropped) would not trigger communicative inferences, because their placement does not seem intentional. Recent research has found that people see behavior as intentional when the outcome causally depends on the agent's desires (Quillien & German, 2021) and when the agent's behavior increases the odds of the outcome happening (Ericson et al., 2023). Integrating these types of processes into the detection of communicative objects is a key step towards having a more flexible framework that can both infer the meaning of communicative objects, and disregard objects that lack a communicative purpose.

Our third assumption was that people can estimate the cost associated with moving objects (for the enforcer) and navigating around them (for the decider), but our model did not explicitly capture how people determine these costs. Recent work shows that, from childhood, people might estimate the effort and difficulty of manipulating objects through an intuitive theory of physical action (Gweon et al., 2017; Yildirim et al., 2019), and integrating this work into our model may enable us to explain communicative objects in more complex situations. This is a direction we hope to pursue in future work.

Our work also has two important limitations. Our first limitation is that we focused our analysis on a specific class of objects where agents reveal their mental states by imposing minimal costs on observers, which we referred to as low-cost communicative blockers (LCCBs). These LCCBs are widespread and easy to find in our everyday lives (see Fig. 1 for examples). Indeed, the presence of a minimal low-cost barrier can even reduce transgressions in children (e.g., children are less likely to peek at the answers in a desk if there is a minimal barrier placed between the two), although this effect even extends to imaginary barriers (Zhao et al., 2020).

How might our framework extend to other types of communicative objects in broader contexts? Most directly, our framework can also explain cases where agents decrease a cost to signal invitation (e.g., leaving a box of cupcakes intentionally open to signal that anyone can grab one, or leaving an office door ajar to indicate we can be interrupted). From our model's perspective, these inferences are symmetrical, with the only difference being that the observer would infer that a communicator intentionally lowered the cost (rather than increased it). This is a prediction that we hope to test in future work.

The most general formulation of our proposal is not intrinsically tied to cost manipulation specifically, but rather to environmental manipulations that reveal mental states. Our work is thus consistent with related work showing that people can use no-cost markers to signal flexible meanings. For instance, when given a sticker to mark which of three cups someone should choose, people use the sticker to signal the right cup. By contrast, when one of the three cups must be avoided, people now use the sticker to indicate avoidance (Misyak et al., 2016). This has been hypothesized to reflect a process known as "virtual bargaining" (Misyak & Chater, 2014; Misyak et al., 2014, 2016), where people produce whichever solution they would have reached if they had gotten the opportunity to have an explicit discussion about it.

More broadly, if people communicate through objects by attempting to elicit mental states in observers, then communicators must also pay attention to additional dimensions that we did not explicitly model. Specifically, for observers to infer mental states from an object's position, the object must be noticeable, and appear intentionally placed. Otherwise, the object might be ignored or dismissed. Intuitively, these features might also play an important role in the use of low-cost communicative blockers (Fig. 1). These additional dimensions are vital for a full computational model of how we infer mental states from objects.

A second key limitation is that our model represents objects in terms of the utilities they provide or impose on agents. This abstraction means that our model does not make any conceptual distinctions between different types of objects as long as they impose the same costs (or provide the same rewards). For instance, our current implementation represents a plastic cup and an expensive water bottle on a table as equivalent (because they impose the same physical cost), even though these two objects would likely elicit different mental-state inferences in observers (in one case, an observer might assume it is trash, while in the other it may be interpreted as a "spot saver"). Similarly, a delivery bag in front of a door may impose a small cost, but knowing what delivery bags are would prevent us from inferring that this is a low-cost communicative blocker. While these examples demonstrate how representing an object's category can strip it of a communicative purpose, there are also cases where representing an object's category strengthens it. For instance, a sisal rope and red velvet rope can both be used to communicate avoidance, but the latter communicates it more strongly (and also carries the implication that there could be negative consequences if the message was ignored, possibly due to knowing that the object was costly and built specifically for this purpose). Intuitively, agents also strategically take this into account when deciding which objects to use to communicate. This is broadly consistent with our account, as it reveals further social reasoning about what observers might find easier to interpret, but is not yet captured by our model.

12.2. Open questions

Our work leaves several major questions open. A first open question stems from our focus on adults: What is the developmental trajectory of how people use communicative objects? Related research shows that the capacities necessary for these kinds of inferences emerge early in development. From early childhood, people can infer the presence of a hidden agent based on the structure of the environment (Keil & Newman, 2015; Ma & Xu, 2013; Newman et al., 2010; Saxe et al., 2005); we can estimate the difficulty associated with fulfilling different tasks (Bennett-Pierre et al., 2018; Gweon et al., 2017; Yildirim et al., 2019); and we can explain behavior in terms of unobservable mental states, like beliefs, desires, and intentions (Gopnik et al., 1997; Wellman, 2014). Most strikingly, recent work has shown that children can also use Theory of Mind to infer the transmission of ideas based on how different agents build similar artifacts (Pesowski et al., 2020; Schachner et al., 2018). This "intuitive archeology" likely shares a common basis with the inferences in our model, opening the possibility that even children can detect and infer the communicative meaning of objects through Theory of Mind. It may even be the case that the inferences we studied here are in fact an extension of people's "intuitive archeology" and ability to reason about the history of objects.

A second open question is about the relative difficulty in creating versus interpreting communicative objects. While our computational model can explain both the creation and interpretation of communicative objects (as revealed in our model simulations), our experimental work focused exclusively on the second component. We did so because, under our account, all people ought to be able to infer the meaning of communicative objects, while the ability to generate them can be limited to a few individuals. We thus do not know the extent to which people can easily create ad-hoc communicative objects in new situations. It is possible that, when creating new communicative objects, people might err on the side of caution and prefer to make the costs higher than would be necessary, to minimize the chance of an agent not realizing the communicative content. As people become more confident that the objects they place are being recognized as communicative, they may subsequently lower the costs. We hope to explore these questions in future work.

A third open question is whether the inferences that we studied here are an extension of pragmatics in language. People's ability to derive non-literal meaning in language is supported by a form of recursive reasoning, captured by the Rational Speech Act (RSA) framework (Frank & Goodman, 2012; Franke & Jäger, 2016; Goodman & Frank, 2016; Scontras et al., 2018). Our model can be thought of as an RSA model where the medium people use to communicate is objects (instead of utterances) and our costs are physical effort (rather than memory retrieval or utterance length). At the same time, the ability to perform recursive social inference is not unique to language and it is possible that the phenomena we studied here reflect a more general, non-linguistic Theory of Mind. As such, these phenomena might constitute a more primitive form of communication that precedes the ability to do pragmatics in language (i.e., a kind of proto-communication). This is a question that goes beyond the scope of our work.

Finally, our work leaves open the question of how to capture other types of inferences that people make from objects. Intuitively, communicative inferences are only a sliver of the social information that we can read from objects. For instance, objects can also lead us to infer aesthetic goals (e.g., placing an object that we like in a visible location within our house), functional goals (e.g., leaving objects like winter gloves next to our front door for convenience), or even personality traits (e.g., inferring that someone is messy based on the general pattern of objects on their desk; Gosling et al., 2002). It is possible that these inferences could be captured through a richer model of Theory of Mind that can consider a broader set of goals that agents have when interacting with the environment, and this is a question that we hope to pursue in future work.

13. Concluding remarks

Humans have a remarkable capacity to share their mental states through their behavior, language, and even the way they arrange objects in their environment. Our work shows one way in which people can share their mental states through objects. And yet, the types of meaning that we give objects is even broader than what we show here—a metal band can signify a lifelong vow, a chiselled stone can commemorate a lost loved one, and a menorah can reveal one's metaphysical beliefs. We hope that our work is a step towards understanding the rich social nature of the physical world.

CRediT authorship contribution statement

Michael Lopez-Brau: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing – original draft, Visualization. **Julian Jara-Ettinger:** Conceptualization, Methodology, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Data availability

Our data/code can be accessed via our project repository: https://osf.io/57n4g.

Acknowledgment

This work was supported by NSF award BCS-2045778 (awarded to Julian Jara-Ettinger).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.cognition.2023.105524.

References

Arechar, A. A., & Rand, D. G. (2021). Turking in the time of COVID. *Behavior Research Methods*, *53*(6), 2591–2595.
Back, E., & Apperly, I. A. (2010). Two sources of evidence on the non-automaticity of true and false belief ascription. *Cognition*, *115*(1), 54–70.
Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, *1*(4), 1–10.
Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition*, *113*(3), 329–349.
Becker, F. D. (1973). Study of spatial markers. *Journal of Personality and Social Psychology*, *26*(3), 439.
Becker, F. D., & Mayo, C. (1971). Measurement effects in studying reactions to spatial invasions. In *Eastern Psychological Assocation Meetings*. ERIC.
Bennett-Pierre, G., Asaba, M., & Gweon, H. (2018). Preschoolers consider expected task difficulty to decide what to do and whom to help. In *Cogsci*.
Birch, S. A., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science*, *18*(5), 382–386.
Conway, B. R., Ratnasingam, S., Jara-Ettinger, J., Futrell, R., & Gibson, E. (2020). Communication efficiency of color naming across languages provides a new framework for the evolution of color terms. *Cognition*, *195*, Article 104086.
Devaine, M., Hollard, G., & Daunizeau, J. (2014). The social Bayesian brain: does mentalizing make a difference when we learn? *PLoS Computational Biology*, *10*(12), Article e1003992.
Edney, J. J., & Jordan-Edney, N. L. (1974). Territorial spacing on a beach. *Sociometry*, 92–104.
Ericson, S. R., Denison, S., Turri, J., & Friedman, O. (2023). Probability and intentional action. *Cognitive Psychology*, *141*, Article 101551.
Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, *336*(6084), 998.
Franke, M., & Jäger, G. (2016). Probabilistic pragmatics, or why Bayes' rule is probably important for pragmatics. *Zeitschrift Für Sprachwissenschaft*, *35*(1), 3–44.
Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naïve theory of rational action. *Trends in Cognitive Sciences*, *7*(7), 287–292.
Gibson, E., Futrell, R., Jara-Ettinger, J., Mahowald, K., Bergen, L., Ratnasingam, S., Gibson, M., Piantadosi, S. T., & Conway, B. R. (2017). Color naming across languages reflects color use. *Proceedings of the National Academy of Sciences*, *114*(40), 10785–10790.
Goodman, N. D., & Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, *20*(11), 818–829.
Gopnik, A., Meltzoff, A. N., & Bryant, P. (1997). *Words, thoughts, and theories*. Cambridge, MA: Mit Press.
Gosling, S. D., Ko, S. J., Mannarelli, T., & Morris, M. E. (2002). A room with a cue: personality judgments based on offices and bedrooms. *Journal of Personality and Social Psychology*, *82*(3), 379.
Griffiths, T. L., Lieder, F., & Goodman, N. D. (2015). Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic. *Topics in Cognitive Science*, *7*(2), 217–229.
Gweon, H., Asaba, M., & Bennett-Pierre, G. (2017). Reverse-engineering the process: Adults' and preschoolers' ability to infer the difficulty of novel tasks. In *CogSci*.
Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83.
Ho, M. K., Littman, M., MacGlashan, J., Cushman, F., & Austerweil, J. L. (2016). Showing versus doing: Teaching by demonstration. *Advances in Neural Information Processing Systems*, 3027–3035.
Hurwitz, E., Brady, T., & Schachner, A. (2019). Detecting social transmission in the design of artifacts via inverse planning. PsyArXiv.
Jacobs, C., Lopez-Brau, M., & Jara-Ettinger, J. (2021). What happened here? Children integrate physical reasoning to infer actions from indirect evidence. *CogSci*.

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

Jara-Ettinger, J., Piantadosi, S., Spelke, E. S., Levy, R., & Gibson, E. (2017). Mastery of the logic of natural numbers is not the result of mastery of counting: Evidence from late counters. *Developmental Science*, *20*(6), Article e12459.

Jara-Ettinger, J., Schulz, L. E., & Tenenbaum, J. B. (2020). The naive utility calculus as a unified, quantitative framework for action understanding. *Cognitive Psychology*, *123*, Article 101334.

Jern, A., Lucas, C. G., & Kemp, C. (2017). People learn other people's preferences through inverse decision-making. *Cognition*, *168*, 46–64.

Keil, F. C., & Newman, G. E. (2015). Order, order everywhere, and only an agent to think: The cognitive compulsion to infer intentional agents. *Mind & Language*, *30*(2), 117–139.

Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science*, *11*(1), 32–38.

Kool, W., & Botvinick, M. (2018). Mental labour. *Nature Human Behaviour*, *2*(12), 899–908.

Lieder, F., & Griffiths, T. L. (2019). Resource-rational analysis: understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, *43*, 1–85.

Liu, S., Ullman, T. D., Tenenbaum, J. B., & Spelke, E. S. (2017). Ten-month-old infants infer the value of goals from the costs of actions. *Science*, *358*(6366), 1038–1041.

Lopez-Brau, M., Kwon, J., & Jara-Ettinger, J. (2022). Social inferences from physical evidence via Bayesian event reconstruction. *Journal of Experimental Psychology: General*.

Lucas, C. G., Griffiths, T. L., Xu, F., Fawcett, C., Gopnik, A., Kushnir, T., Markson, L., & Hu, J. (2014). The child as econometrician: A rational model of preference understanding in children. *PLoS One*, *9*(3), Article e92160.

Ma, L., & Xu, F. (2013). Preverbal infants infer intentional agents from the perception of regularity. *Developmental Psychology*, *49*(7), 1330–1337.

Matsumoto, D., & Yoo, S. H. (2006). Toward a new generation of cross-cultural research. *Perspectives on Psychological Science*, *1*(3), 234–250.

McDermott, J. H., Schultz, A. F., Undurraga, E. A., & Godoy, R. A. (2016). Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature*, *535*(7613), 547–550.

Misyak, J. B., & Chater, N. (2014). Virtual bargaining: A theory of social decision-making. *Philosophical Transactions of the Royal Society, Series B (Biological Sciences)*, *369*(1655), Article 20130487. http://dx.doi.org/10.1098/rstb.2013.0487.

Misyak, J. B., Melkonyan, T., Zeitoun, H., & Chater, N. (2014). Unwritten rules: virtual bargaining underpins social interaction, culture, and society. *Trends in Cognitive Sciences*, *18*(10), 512–519.

Misyak, J. B., Noguchi, T., & Chater, N. (2016). Instantaneous conventions: The emergence of flexible communicative signals. *Psychological Science*, *27*(12), 1550–1561.

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.

Oey, L. A., Schachner, A., & Vul, E. (2022). Designing and detecting lies by reasoning about other agents. *Journal of Experimental Psychology: General*.

Pelz, M., Schulz, L., & Jara-Ettinger, J. (2020). The signature of all things: Children infer knowledge states from static images. PsyArXiv.

Pesowski, M. L., Kelemen, D., & Schachner, A. (2021). Children use artifacts to infer others' shared interests. In *Proceedings of the Annual Meeting of the Cognitive Science Society. Vol. 43*.

Pesowski, M. L., Quy, A. D., Lee, M., & Schachner, A. (2020). Children use inverse planning to detect social transmission in design of artifacts. In *Proceedings of the Annual Conference of the Cognitive Science Society*.

Piantadosi, S. T., Jara-Ettinger, J., & Gibson, E. (2014). Children's learning of number words in an indigenous farming-foraging group. *Developmental Science*, *17*(4), 553–563.

Poortinga, Y. H. (1989). Equivalence of cross-cultural data: An overview of basic issues. *International Journal of Psychology*, *24*(6), 737–756.

Powell, L. J. (2022). Adopted utility calculus: Origins of a concept of social affiliation. *Perspectives on Psychological Science*, *17*(5), 1215–1233.

Quillien, T., & German, T. C. (2021). A simple definition of 'intentionally'. *Cognition*, *214*, Article 104806.

Rand, D. G. (2016). Cooperation, fast and slow: Meta-analytic evidence for a theory of social heuristics and self-interested deliberation. *Psychological Science*, *27*(9), 1192–1206.

Rossano, F., Fiedler, L., & Tomasello, M. (2015). Preschoolers' understanding of the role of communication and cooperation in establishing property rights. *Developmental Psychology*, *51*(2), 176.

Royka, A., Chen, A., Aboody, R., Huanca, T., & Jara-Ettinger, J. (2022). People infer communicative action through an expectation for efficient communication. *Nature Communications*.

Saxe, R., Tenenbaum, J., & Carey, S. (2005). Secret agents: Inferences about hidden causes by 10- and 12-month-old infants. *Psychological Science*, *16*(12), 995–1001.

Schachner, A., Brady, T., Oro, K., & Lee, M. (2018). Intuitive archeology: Detecting social transmission in the design of artifacts. PsyArXiv.

Scontras, G., Tessler, M. H., & Franke, M. (2018). Probabilistic language understanding: An introduction to the Rational Speech Act framework. URL https://gscontras.github.io/problang.

Shaffer, D. R., & Sadowski, C. (1975). This table is mine: Respect for marked barroom tables as a function of gender of spatial marker and desirability of locale. *Sociometry*, 408–419.

Shafto, P., Goodman, N. D., & Griffiths, T. L. (2014). A rational account of pedagogical reasoning: Teaching by, and learning from, examples. *Cognitive Psychology*, *71*, 55–89.

Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., & Botvinick, M. M. (2017). Toward a rational and mechanistic account of mental effort. *Annual Review of Neuroscience*, *40*, 99–124.

Sommer, R., & Becker, F. D. (1969). Territorial defense and the good neighbor. *Journal of Personality and Social Psychology*, *11*(2), 85.

Ullman, T., Baker, C., Macindoe, O., Evans, O., Goodman, N., & Tenenbaum, J. B. (2009). Help or hinder: Bayesian models of social goal inference. In *Advances in Neural Information Processing Systems* (pp. 1874–1882).

Van de Vijver, F. J., & Leung, K. (2021). *Methods and Data Analysis for Cross-Cultural Research. Vol. 116*. Cambridge University Press.

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science*, *311*(5765), 1301–1303.

Wellman, H. M. (2014). *Making Minds: How Theory of Mind Develops*. Oxford University Press.

Yildirim, I., Saeed, B., Bennett-Pierre, G., Gerstenberg, T., Tenenbaum, J., & Gweon, H. (2019). Explaining intuitive difficulty judgments by modeling physical effort and risk. arXiv preprint arXiv:1905.04445.

Zhao, L., Zheng, Y., Compton, B. J., Qin, W., Zheng, J., Fu, G., Lee, K., & Heyman, G. D. (2020). The moral barrier effect: Real and imagined barriers can reduce cheating. *Proceedings of the National Academy of Sciences*, *117*(32), 19101–19107.