What happened here? Children integrate physical reasoning to infer actions from indirect evidence

Colin Jacobs, Michael Lopez-Brau, & Julian Jara-Ettinger
{colin.jacobs, michael.lopez-brau, julian.jara-ettinger}@yale.edu
Department of Psychology, Yale University New Haven, CT 06511 USA

Abstract

As we navigate through the world, we often leave traces of our actions: a broken branch, a footprint in the mud, a dirty coffee mug at a desk. As observers, these traces enable us to make surprisingly complex social inferences about the actions that may have caused them: what the other person may have been doing, what their likely goals were, and more. But how might a conspicuous lack of evidence prompt similar reasoning? We hypothesize that children consider the presence and absence of physical evidence to infer possible prior actions and their outcomes. To test this hypothesis, we ask children to infer which of two bowls (each containing different materials) was acted upon without witnessing the action directly. In support of this proposal, we found that children readily reconstruct an agent's actions after observing indirect evidence. Importantly, they are also able to use the difficulty of concealing such evidence to interpret its absence.

Keywords: Theory of Mind; Social Cognition; Cognitive Development.

Introduction

Humans have a remarkable capacity to understand and navigate the social world. Integral to our social skills is the ability to represent other people's behavior in terms of mental states such as beliefs, desires, and intentions (Gopnik & Meltzoff, 1998; Wellman, 2014; Baker, Jara-Ettinger, Saxe, & Tenenbaum, 2017). These mental-state attributions enable us to explain other people's actions (Malle, 1999); they guide how we interact with them (Repacholi & Gopnik, 1997); they allow us to predict how others might act (Liu, Ullman, Tenenbaum, & Spelke, 2017; Jara-Ettinger, Schulz, & Tenenbaum, 2020); and they form the basis of many social activities, from language understanding to moral reasoning (Young, Cushman, Hauser, & Saxe, 2007; Sperber & Wilson, 1986).

At the heart of this capacity, called a *Theory of Mind*, is the ability to infer other people's mental states by watching how they behave. From infancy, this capacity is structured around an expectation that agents act to minimize the costs that they incur and maximize the rewards that they obtain (Gergely & Csibra, 2003; Jara-Ettinger, et al., 2016; Liu et al., 2017). Through this assumption, people's actions reveal their goals, preferences, beliefs, and desires (see Jara-Ettinger et al., 2016 for review).

Human social cognition, however, goes beyond an ability to infer mental states from direct, observable action. Imagine that you're going out of town for the weekend and you ask your housemates not to let anyone sleep in your room while you're away. If you returned to find a disheveled bed, you could reasonably infer that someone stayed over despite your request. Conversely, if you returned to find a perfectly-made bed, as you had left it, you could assume that no one used your room while you were gone. But what if you heard from mutual friends about a party they threw while you were away? In this case, the lack of evidence that someone slept in your room may not feel as convincing. Did they make the bed afterwards to cover up their transgression? How feasible would it be for them to perfectly conceal the evidence that someone slept there? If you had a peculiar way of tucking in your sheets, what are the chances that your roommates would've known to do that?

A growing body of work has found that, given indirect physical evidence, such as breadcrumbs on the floor or a messy bedroom, people can make a range social inferences, including what people did, their goals, their personalities, and even the transmission of ideas across people (Goslin, Ko, Mannarelli, & Morris, 2002; Hurwitz, Brady, & Schachner, 2019; Lopez-Brau, Kwon, & Jara-Ettinger, 2020). Moreover, recent research suggests that this capacity emerges early in childhood, with preschoolers already being able to infer mental states from indirect traces in the environment (Pelz, Schulz, & Jara-Ettinger, 2020; Pesowski, Quy, Lee, & Schachner, 2020).

This research has largely focused on how we infer agents' behavior based on the goals that they were pursuing and the traces that they leave in their environment. However, as the examples above show, social inferences from indirect evidence also rely on our understanding of what types of actions leave observable traces, as well as when and how these traces can be removed. Here we propose that, from childhood, social inferences from indirect physical evidence depend not only on our expectations about how agents act, but also on the likelihood that agents' actions would leave observable environmental traces (as well as the relative ease or difficulty of removing these traces). This capacity is provided by our intuitive physics—our mental models that allow us to understand how objects and forces interact with each other (Battaglia, Hamrick, & Tenenbaum, 2013). Under this view, children have a causal understanding of how agents' actions leave traces in the environment, as determined by their intuitive physics, and they use this understanding to work backwards from a physical scene to determine what might have happened before.

While a growing set of studies have shown that people can infer mental states and actions from physical evidence, to our knowledge, no work has tested whether this early-emerging capacity supports inferences not only from direct environmental traces, but also the absence of traces. Here we test if children can reconstruct prior actions of an agent given indirect physical evidence. To test this possibility, we adapted an experimental paradigm used to test deception in children (Evans, Xu, & Lee, 2011). In that paradigm, children, alone in a room, can lift a cup to find what's inside. When they do this, however, the cup spills rice out, leaving a clear trace. This paradigm has been classically used to test whether children engage in deception (denying that they lifted the cup) when the lie cannot be justified. Here, we modified this paradigm so that different bowls leave different traces which are harder or easier to clean up. We then presented children with a scene containing different types of (or lack of) physical evidence, and asked them to infer what another agent did without ever witnessing the action or an outcome that matches the final scene.

We focus our study on five- and six-year-olds. Prior studies have found that by age four children can use physical traces for social inference (Pelz et al., 2020, Pesowski et al., 2020), and may engage in counterfactual reasoning under specific conditions (Nyhout & Ganea, 2020), however; in these tasks, children are explicitly presented with possible alternative outcomes (either visually displayed or verbally prompted during the test question). In contrast, for success in our task, children must use a single static scene to reason about the evidence lifting either bowl is likely to have left, the difficulty of concealing that evidence, and how these factors reveal prior action given the observed outcome. We also expected this task to be more difficult for younger ages which may struggle with more complex counterfactual reasoning (McCormack et al., 2018). Thus, we chose five- and six-year-olds to account for the increased difficulty of our task while still investigating potential developmental differences.

Experiment

Children were introduced to two upside-down bowls, one concealing rocks and the other concealing rice underneath. After learning the contents of each bowl, children saw a small pile of rocks, a small pile of rice, or no physical evidence between the two bowls, and they were asked which of the two bowls an agent had lifted.

If children are able to use physical traces to infer an agent's past behavior, then observing either type of material in the final scene should reveal that the agent acted upon the corresponding bowl. Additionally, when no evidence is present, children should infer that the agent interacted with the bowl containing the material that was easiest to conceal—the rock-filled bowl. If children are simply matching type of material to bowl to identify prior actions, we expect them to be at chance in this no-evidence condition. If, however, five- to six-year-olds interpret the lack of evidence by considering potential actions and their expected outcomes, they may successfully identify which bowl the agent lifted. All materials, procedures, predictions, exclusion criteria and analyses were pre-registered.

Method

Participants 108 five- to six-year-olds (mean age: 5.94 years, range: 5.01 - 6.99 years) were recruited and tested online via Zoom. Five additional participants were recruited but not included because they failed to answer inclusion questions correctly (n = 2), because they declined to complete the study (n = 1), because of technical issues with the Zoom call (n = 1), and because of family interference (n = 1).

Participants were randomly assigned to one of two conditions, either "evidence" or "no-evidence." Note that both types of evidence (rocks and rice) were run as a single condition counterbalanced across participants, but are presented below separately for clarity. Sample sizes were determined using a power analysis with effect-sizes estimated from pilot data (68.4% success, n = 38 in the no-evidence condition; 100% success, n = 15 in the evidence condition), aimed at having power > 0.9 in both conditions. Our analyses resulted in a sample size of n = 72 for the no-evidence condition, and n = 36 for the evidence condition.

Stimuli Stimuli consisted of a blue and an orange bowl, a pile of rocks, a pile of rice, and a male puppet. These stimuli were used to create two warm-up videos where an experimenter lifted each bowl to reveal what as underneath, and three test images displaying the two bowls with either no evidence, or observable evidence (rice or rocks). The videos and pictures were produced multiple times to counter-balance the contents and position of the bowls.

Procedure Participants were first introduced to the colored bowls, one orange and one blue. They were told that these are "very special bowls. They each have something secret inside of them," and then were asked to select one to peek underneath first.

Children then watched a video of the selected bowl being lifted. While the bowl was lifted and the materials spilled out, the experimenter exclaimed "Look! There are [rocks/rice] underneath this bowl! Let's try and put it back the way it was." The video then showed a hand attempt to push the materials back into a neat pile. As this occurred, the experimenter noted the difficulty of cleaning up verbally (i.e., "Hmm, it's really hard to put the rice back underneath the bowl," or "It's easy to put the rocks back underneath the bowl."). The bowl was then placed back into its original position, but a minimal amount of the material remained unconcealed (Figure 1a). The experimenter then repeated this process with the second bowl. Note that though the difficulty of cleaning each material varied, a similar amount of rocks and rice are left outside of the bowls once they are put back the original positions (i.e. the rice was more difficult to clean, but this did not correspond to more rice being left outside of the bowl during this demonstration). Thus, by the end of the warmup, participants knew the contents of each bowl; they knew that the contents spill out when the bowls were lifted;

and they saw the agent attempt to conceal the contents, failing to fully conceal the evidence for both bowls.

Participants were next presented with a new clean setup using the same two bowls, but with no rice or rocks visible in the scene. The experimenter introduced participants to a male puppet ("Michael") and explained that the puppet did not know what was under the bowl, and was only told that "it was a secret, and he was not allowed to peek." The experimenter then explained that the puppet was left alone with the bowls, and explained that "even though I told Michael not to peek, when I came back, this is what I found," as they displayed a test image (depending on condition, see Figure 1b). In the evidence condition, either rocks or rice were visible directly in between the two bowls, with the type of evidence counterbalanced across participants (Figure 1b). The size of the evidence was matched in both cases. In the no-evidence condition, the test image was identical to the initial image, with no physical traces of rocks or rice visible (Figure 1b).

When the test image was presented, participants were told that the puppet had admitted to "peeking under just one of the bowls, but he wouldn't tell me which one!" Participants were then asked "Can you help me? Can you tell me which bowl Michael peeked under?" After the test question, participants were asked to recall which bowl had rocks and which bowl had rice as inclusion criteria. The side that each color bowl was on, materials contained by either color, and evidence type in the evidence condition were counterbalanced across participants.

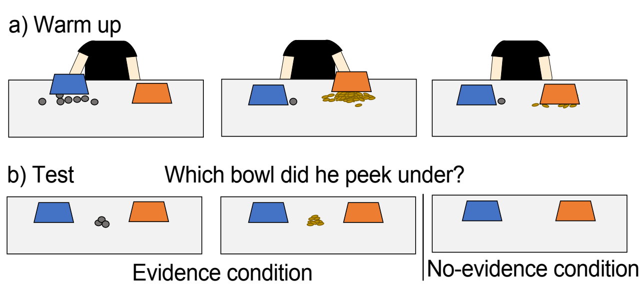


Figure 1: Procedure of Experiment 1. a) Warm-up phase. The experimenter lifted each bowl, the material contained spilled out, and the experimenter attempted to put it back under the bowl, failing to fully conceal the evidence with both bowls. b) Schematic of test trials. Participants in the evidence condition either saw a pile of rocks or rice between the two cups (counter-balanced across participants), and participants in the no-evidence condition saw the two bowls with no visible traces of rocks or rice.

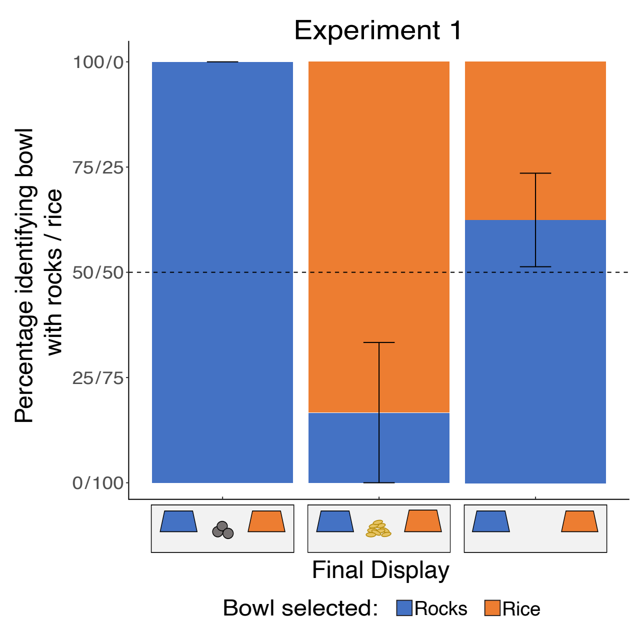
Results

We first confirmed that participant's choice for which bowl to view was not biased in a way that could influence our results. 57.4% (n =62) participants selected to the blue bowl, and 42.6% (n = 46) selected the orange bowl. Due to the counterbalanced association between bowl and contents, 50% (n = 54) first saw rocks, and 50% (n = 54) first saw rice.

As pre-registered, a coder blind to the participant's final response coded for experimenter error, family interference, or substantial distraction during the testing session. This coding was done using videos (n = 91) or experimenter notes when permission to record was not given (n = 17). One participant was excluded and replaced due to family interference.

In the evidence condition, 91.67% (n = 33; 95% CI: 83.33-100)[1] of participants chose the bowl which contained the evidence visible in the final scene, and children's success was comparable for both types of evidence (type counterbalanced across trials). When rocks were shown, 100% (n = 18) of participants selected the rock bowl, and when rice were shown, 83.33% (n = 15; 95% CI: 66.67-100) selected the rice bowl (Figure 2). Critically, the evidence always appeared between the two bowls and was matched in size. Children could therefore not succeed in this task using spatial or magnitude cues. Nonetheless, success in the evidence condition could still be explained by a superficial association between the evidence and the contents inside each bowl. If children do rely on physical reasoning, they should also succeed in the no evidence condition.

In the no-evidence condition, 62.50% (n = 45; 95% CI: 51.39-73.61) of participants chose the bowl containing rocks. To see if age influenced children's success in the no-evidence condition, we ran a logistic regression predicting participant choice (coded as correct or incorrect) as a function of age (as a continuous variable). This analysis revealed no evidence of developmental change ($\beta = 0.298$; $p = 0.48$), suggesting that the weaker effect size in the no-evidence condition did not stem from a developmental change in our age group.



[1] Due to conceptual limitations associated with NHST (Cohen, 1994; Cumming, 2014), we rely on confidence intervals over effect sizes as our main method our analyses. We consider a result to be statistically reliable whenever the bootstrapped 95% confidence interval does not include chance performance.

Figure 2: Experiment results. The x axis shows condition (with evidence type presented separately in the evidence condition), and the y axis shows the percentage of participants who identified each bowl as the one which was acted upon. Vertical lines show 95% bootstrapped confidence intervals and the horizontal dotted line represents chance.

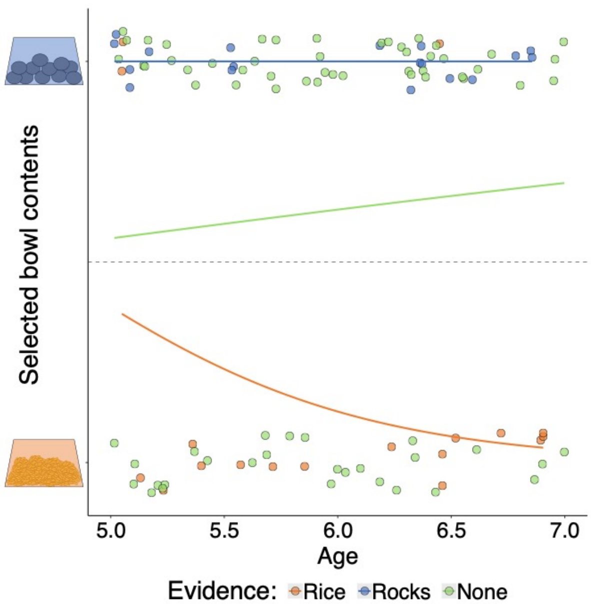


Figure 3: Participant responses to each type of evidence presented. Each dot represents a participant's answer. The x-axis shows their age, the y-axis shows which bowl they identified during the test question. Data is jittered slightly on the y-axis for visibility. Color indicates the evidence presented during the final slide. Each line shows a logistic regression.[2]

Discussion

From childhood, human social cognition affects not only how we interpret other people's behavior, but also how we make sense of the physical traces that they leave in the environment (Goslin et al., 2002; Hurwitz et al., 2019; Lopez-Brau et al., 2020; Pelz et al., 2020; Pesowski et al., 2020). Our work shows that, by age five, children's ability to infer what an agent did based on indirect physical evidence relies on their understanding of how actions leave observable traces in the environment. In the evidence condition, participants successfully inferred that an agent had lifted whichever cup contained the observable contents, even when the material appeared spatially between the two bowls. While children in this condition could succeed simply by associating the observable evidence with the corresponding bowl, children were also able to infer what an agent did when no evidence was visible. In the no-evidence condition, participants reliably inferred that the agent had lifted the bowl with rocks.

Children's success in this condition is more challenging to explain via superficial associations. During the warm-up phase, the experimenter failed to hide the same amount of evidence for both bowls, always leaving a few rocks or a bit of rice visible outside of the bowl. Thus, children so far had only seen no-evidence images in the initial setup where no one had interacted with either bowl. This suggests that children solved the task by considering the relative difficulty of hiding rice and rocks under the bowls. In particular, the results indicate that children rely on physical reasoning to infer what kinds of traces different actions leave behind, as well as which traces are easier to conceal.

Further evidence that children did not rely on a simple superficial association comes from the fact that children's performance on the no-evidence condition was weaker relative to the evidence condition. If children relied on simple cues in all trials, the effect sizes should be comparable. From an intuitive physics standpoint, however, the weaker effect size is consistent with the additional cognitive demands required to solve the no-evidence trial. Here, children must consider not only the effect of lifting each bowl, but also the relative difficulty of removing the evidence.

One open question left by this study is whether children spontaneously attend to the physical information and infer the relative ease or difficulty of different actions. In our task, the experimenter showed how lifting the bowl caused its contents to spill out, and noted whether putting the contents back underneath the bowl was easy or hard. Thus, we do not know whether children would spontaneously infer that rice is harder to conceal than rocks due to its physical properties alone (without demonstrations). Similarly, we do not know to what extent children rely on the verbal descriptions offered by the experimenter to reason about the difficulty of cleaning each material. Though to succeed in our task children must consider physical evidence to infer a prior action, it's possible that children may not spontaneously attend to the difficulty of concealing each material without these descriptions. As such, our results suggest that children can reason about different potential actions and their expected physical traces to infer what actions an agent took, but do not reveal the nuance of when they may naturalistically apply this ability.

Additionally, our study found no developmental change in the age range that we considered. We therefore do not know when this capacity emerges, and it is possible that it could be at work before age five. Note, however, that our task requires children to integrate multiple capacities: inferences about how actions cause observable traces, an understanding of the difficulty of different actions, and an ability to consider multiple hypothetical alternatives. Thus, it is possible that younger children might fail in this task due to the lack of development of any of these different capacities. Isolating the source of each capacity and its development is a direction that we hope to pursue in future work.

[2] As suggested by reviewers, we conducted an exploratory logistic regression for the evidence condition when rice was visible in the final scene. This analysis revealed no significant effect of age on success within this subset of the evidence condition (β=1.469, p =0.214).

Finally, our work focused on the role of physical reasoning in children's ability to infer what actions an agent took with no information about the agent's desires or preferences. Therefore, while children are reasoning about the actions of agents, our task does not directly require Theory of Mind to succeed. In more complex situations, both knowledge about what agents prefer and physical reasoning must be integrated to draw accurate inferences. Our work leaves open the question whether children can jointly combine information about how agents' mental states lead to different actions, which in turn leave physical traces.

Overall, our work shows that children can make rich social inferences from physical observations. Moreover, this capacity is not limited to linking physical traces to the corresponding actions, and it extends to inferences about the absence of evidence. Our findings suggest that, from a young age, children rely their causal model of the physical world to derive social inferences from physical scenes, highlighting the flexibility of social reasoning in early childhood.

References

Asaba, M., Ong, D. C., & Gweon, H. (2019). Integrating expectations and outcomes: Preschoolers' developing ability to reason about others' emotions. *Developmental psychology*.

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in cognitive sciences*, *14*(3), 110-118.

Battaglia, P. W., Hamrick, J. B., & Tenenbaum, J. B. (2013). Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences*, *110*(45), 18327-18332.

Cohen, J. (1994). The world is round (p<. 05). *American Psychologist*, *49*(12), 997-1003.

Cumming, G. (2014). The new statistics: Why and how. *Psychological science*, *25*(1), 7-29.

Evans, A. D., Xu, F., & Lee, K. (2011). When all signs point to you: Lies told in the face of evidence. *Developmental Psychology*, *47*(1), 39.

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naıve theory of rational action. *Trends in cognitive sciences*, *7*(7), 287-292.

Gopnik, A., & Meltzoff, A. N. (1998). *Words, Thoughts, and Theories,* 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.

Harris, P. L., Johnson, C. N., Hutton, D., Andrews, G., & Cooke, T. (1989). Young children's theory of mind and emotion. *Cognition & Emotion*, *3*(4), 379-400.

Hurwitz, E., Brady, T., & Schachner, A. (2019). *Detecting social transmission in the design of artifacts via inverse planning.* PsyArXiv. https://doi.org/10.31234/osf.io/qup6b

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

Kubricht, J. R., Holyoak, K. J., & Lu, H. (2017). Intuitive physics: Current research and controversies. *Trends in cognitive sciences*, *21*(10), 749-759.

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. (2020). Mental state inference from indirect evidence through Bayesian event reconstruction. *Proceedings of the 42nd Annual Conference of the Cognitive Science Society*

Malle, B. F. (1999). How people explain behavior: A new theoretical framework. *Personality and social psychology review*, *3*(1), 23-48.

McCormack, T., Ho, M., Gribben, C., O'Connor, E., & Hoerl, C. (2018). The development of counterfactual reasoning about doubly-determined events. *Cognitive Development*, *45*, 1-9.

Nyhout, A., & Ganea, P. A. (2019). Mature counterfactual reasoning in 4-and 5-year-olds. *Cognition*, *183*, 57-66.

Pelz, M., Schulz, L., & Jara-Ettinger, J. (2020). The Signature of All Things: Children Infer Knowledge States from Static Images. *Proceedings of the 42nd Annual Conference of the Cognitive Science Society.*

Pesowski, M., Quy, A., Lee, M., & Schachner, A. (2020). *Children use inverse planning to detect social transmission in design of artifacts.* PsyArXiv. https://doi.org/10.31234/osf.io/my6ba.

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: evidence from 14-and 18-month-olds. *Developmental psychology*, *33*(1), 12.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition* (Vol. 142). Cambridge, MA: Harvard University Press

Young, L., Cushman, F., Hauser, M., & Saxe, R. (2007). The neural basis of the interaction between theory of mind and moral judgment. *Proceedings of the National Academy of Sciences*, *104*(20), 8235-8240.