Infancy

WILEY

INFANCY THE OFFICIAL JOURNAL OF THE INTERNATIONAL CONGRESS OF INFANT STUDIES

RESEARCH ARTICLE

15-Month-Olds' Understanding of Imitation in Social and Instrumental Contexts

Shannon Yasuda[1] | Wenjie Li[2] | Deisy Martinez[1] | Brenden M. Lake[1,2] | Moira R. Dillon[1]

[1]Department of Psychology, New York University, New York, New York, USA | [2]Center for Data Science, New York University, New York, New York, USA

Correspondence: Shannon Yasuda (sy3494@nyu.edu) | Moira R. Dillon (moira.dillon@nyu.edu)

Received: 28 June 2024 | **Revised:** 16 November 2024 | **Accepted:** 6 January 2025

Funding: This work was supported by a National Science Foundation CAREER Award (DRL1845924 to MRD) and a DARPA grant on Machine Common Sense (HR001119S0005 to BML and MRD).

Keywords: action understanding | affiliation | imitation | infancy | social cognition

ABSTRACT

From early in development, humans use imitation to express social engagement, to understand social affiliations, and to learn from others. Nevertheless, the social and instrumental goals that drive imitation in everyday and pedagogical contexts are highly intertwined. What cues might infants use to infer that a social goal is driving imitation? Here we use minimal and tightly controlled visual displays to evaluate 15-month-olds' attribution of social goals to imitation. In particular, we ask whether they see the very same simple, imitative actions shared between two agents as social or nonsocial when those actions occur in the absence or presence of intentional cues such as obstacles, object goals, and efficient, causally effective action. Our results suggest that infants' attributing social value to imitation only in the absence of such intentional cues may be a signature of humans' early understanding of imitation. We propose, moreover, that a systematic evaluation of a set of simple scenarios that probe candidate principles of early knowledge about social and instrumental actions and goals is possible and promises to inform our understanding of the foundational knowledge on which human social learning is built, as well as to aid the building of human-like artificial intelligence.

1 | Introduction

Imitation of the expressions, goals, and actions of others is foundational to human social learning (e.g., Gergely and Csibra 2020; Legare and Nielsen 2015; Tomasello et al. 2005). What kinds of imitative actions do infants and children see as social and so as an opportunity for learning? Young infants themselves imitate the facial gestures and expressions of other individuals with whom they are engaging socially (Meltzoff and Moore 1999) and look longer at animated characters who imitate versus do not imitate other animated characters (Powell and Spelke 2018b). Young infants also show consistent third-party predictions about the social value of others' novel but imitative sounds and actions (Powell and Spelke 2018a) when judging both social group membership (Powell and Spelke 2013) and dyadic relationships (Powell and Spelke 2018a).

Young infants' first-person preferences and third-person predictions about the social value of imitation have been found in scenarios without environmental obstacles, constraints to characters' actions, or objects or locations that were the goals of those actions. Adults' judgments about the potential social value of such unconstrained and nonintentional actions are consistent with these findings with infants. For example, when adults see a single animated character produce a series of actions in the absence of any causal outcomes to objects, they think that the character's movement itself is the goal, and they may look for a social explanation for the character's actions, for example, the character is "dancing." But when the character makes the very same series of actions to avoid obstacles and interact with objects, adults instead infer that the character's goal is object directed, for example, the character is gathering and rearranging the objects (Schachner and Carey 2013). How, then, might infants' social evaluation of imitative actions be affected by the presence of object goals?

On the one hand, even young infants are keenly sensitive to obstacles, efficiency, and causal efficacy when interpreting other agents' goal-directed actions on objects. Young infants predict, for example: that agents will take a new but efficient path to a goal object when an obstacle originally blocking a goal object is removed (Csibra et al. 1999; Gergely et al. 1995; Skerry, Carey, and Spelke 2013; Stojnić et al. 2023); that agents will show consistent object-based goals (Stojnić et al. 2023; Woo, Liu, and Spelke 2024; Woodward 1998); and that agents' actions on objects are causally effective (Liu, Brooks, and Spelke 2019; Muentener and Carey 2010). On the other hand, young infants seem insensitive to the social value of imitative actions on objects. For example, older but not younger infants look longer and smile at an individual who has imitated their own actions on an object versus someone who has not (Agnetta and Rochat 2004; Carpenter, Uebel, and Tomasello 2013; Meltzoff 1988; Meltzoff and Moore 1999). And, older infants imitate the actions of another individual who has looked at them and then acted unusually to effect change on an object, like turning a light on with their forehead instead of with their hands (Gergely, Bekkering, and Király 2002; Meltzoff 1988).

Toddlers in the second year of life thus appear to become ever more attuned to the potential social value of simultaneously imitative and object-directed actions, further shown by their tendency to over imitate a social partner's actions on objects (e.g., Buttelmann et al. 2013) and appreciate "ritualistic" actions on objects by members of the same social group (e.g., Liberman, Kinzler, and Woodward 2018). In particular, and consistent with the suggestion that older infants and toddlers are rational and flexible in their inferences about the social value of object-directed imitation (e.g., Gergely, Bekkering, and Király 2002), older infants expect two individuals to affiliate after they have performed the same, inefficient action on an object (e.g., turning on a light with their head), but infants do not expect two individuals to affiliate if their same actions are efficient relative to an environmental constraint (e.g., turning on a light with their head when their arms were wrapped in a blanket; Liberman, Kinzler, and Woodward 2018). Similarly, older infants expect two individuals who chose to wear the same clothing (e.g., containing a pocket) to affiliate, but only when their clothing choices do not reflect an instrumental goal (e.g., storing toys in their pocket; Bian and Baillergeon 2022).

Nevertheless, the intentional and social factors in everyday and pedagogical instances of imitation are highly intertwined. And so, in the present work we aim to identify the role of a specific set of intentional cues—obstacles, object goals, and efficient, causally effective actions—on older infants' predictions about the social value of imitation. To do so, we employ a minimal, tightly controlled social context with minimally contrastive scenarios. Our stimuli and design are adapted from previous work that evaluated the predictions of both 11-month-old infants and learning-driven neural-network models about the actions, goals, and rationality of single, minimal agents who interacted with obstacles and objects in a grid-world environment (Stojnić et al. 2023). Here, we use this paradigm to evaluate whether 15-month-olds think the very same simple, imitative actions shared between two minimal agents are social when they occur either in the absence or in the presence of multiple intentional cues. We find that infants attribute social value to imitation only in the absence of such intentional cues, and we suggest that this kind of attribution may be a signature of humans' early understanding of imitation.

2 | Experiment 1

2.1 | Methods

2.1.1 | Participants

A simulation-based power analysis for mixed-model linear regressions (Green and MacLeod 2016) based on two prior, related tasks (Stojnić et al. 2023) suggested that a sample size of $N = 32$ per task would lead to 81% power to determine whether infants have differential expectations on the current tasks. Considering this and to accommodate the tasks' counterbalancing, we planned a sample size of $N = 32$ per task.

Forty-seven typically developing 15-month-olds (Imitation task: $N = 33$, *M*age: 14.86 m, range: 14.50 m–15.36 m, 15 girls; Instrumental task: $N = 32$, *M*age: 14.93 m, range: 14.53 m–15.49 m, 12 girls) born at ≥ 37 weeks gestational age participated. Participants were asked to complete both an Imitation task and an Instrumental task during different Zoom sessions on different days within 2 weeks. Data collection continued until enough infants were run to include at least $N = 32$ in each task. The Instrumental task having reached that number, the last infant for the Imitation task contributed data to both tasks, resulting in $N = 33$ sessions in the Imitation task and $N = 32$ sessions in the Instrumental task. Half of the infants completed the Imitation task first, and $N = 17$ infants contributed data to both tasks. An additional 34 sessions (Imitation task: 15; Instrumental task: 22) were excluded based on predetermined criteria, including: not completing the session (8); technical failure (9); poor video quality and/or missing video (6); caretaker interference (8); and looking time < 1 s to either test trial or < 2 s to two or more familiarization trials (3). Three more sessions were excluded post hoc for extreme values (> 50 s) to one test trial, identified through Cook's Distance. All determinations of exclusion were made by an experimenter masked to infants' performance. Informed consent was obtained prior to each session and the use of human participants for this study was approved by the Institutional Review Board at New York University.

2.1.2 | Materials

The stimuli for the Imitation and Instrumental tasks were procedurally generated with code adapted from Stojnić et al. (2023). Short silent animated videos presented simple shapes as agents without eyes or limbs undertaking basic movements in a 2D grid world shown from above (Figure 1). Each task included eight familiarization videos, which varied slightly in the starting location of the two target agents across trials, followed by two test videos, each presenting a different outcome.

Inspired by Powell and Spelke (2018a), the Imitation task tested whether infants predicted that an agent would be more likely to approach another agent whose actions it imitated over an agent whose actions it did not imitate (Figure 1A). In four familiarization trials, one agent performed a simple action, and then the imitator performed the same action. In the other four familiarization trials, a third agent performed a different action, after which the imitator continued to perform its original action. These two kinds of familiarization trials alternated during the familiarization phase. All three agents were present in each video as was an additional shape, but none of the agents interacted with this last shape nor did it move. In the two test videos, only the three agents were present. In the same-action test video, the imitator approached the agent it had imitated during familiarization (expected), but in the different-action test video, the imitator approached the agent it had not imitated during familiarization (unexpected). If infants understood the shapes' actions as agentive and social and if they expected the imitating agent to approach the agent it had imitated, then they would look longer to the different-action/unexpected test video.

The Instrumental task was similar to the Imitation task but had one critical difference: When the imitator moved the same way as one of the other two agents, its motion included efficient action around a black obstacle in the grid world to contact the fourth, stationary shape and change its color. Because the imitator's actions were rational (Gergely et al. 1995) and causally effective (Liu, Brooks, and Spelke 2019), they could be seen as directed toward an object, not a social, goal (Schachner and Carey 2013). If infants' predictions about the imitator's subsequent approach to one of the two agents were specific to contexts in which there were no obstacles or goal objects present,

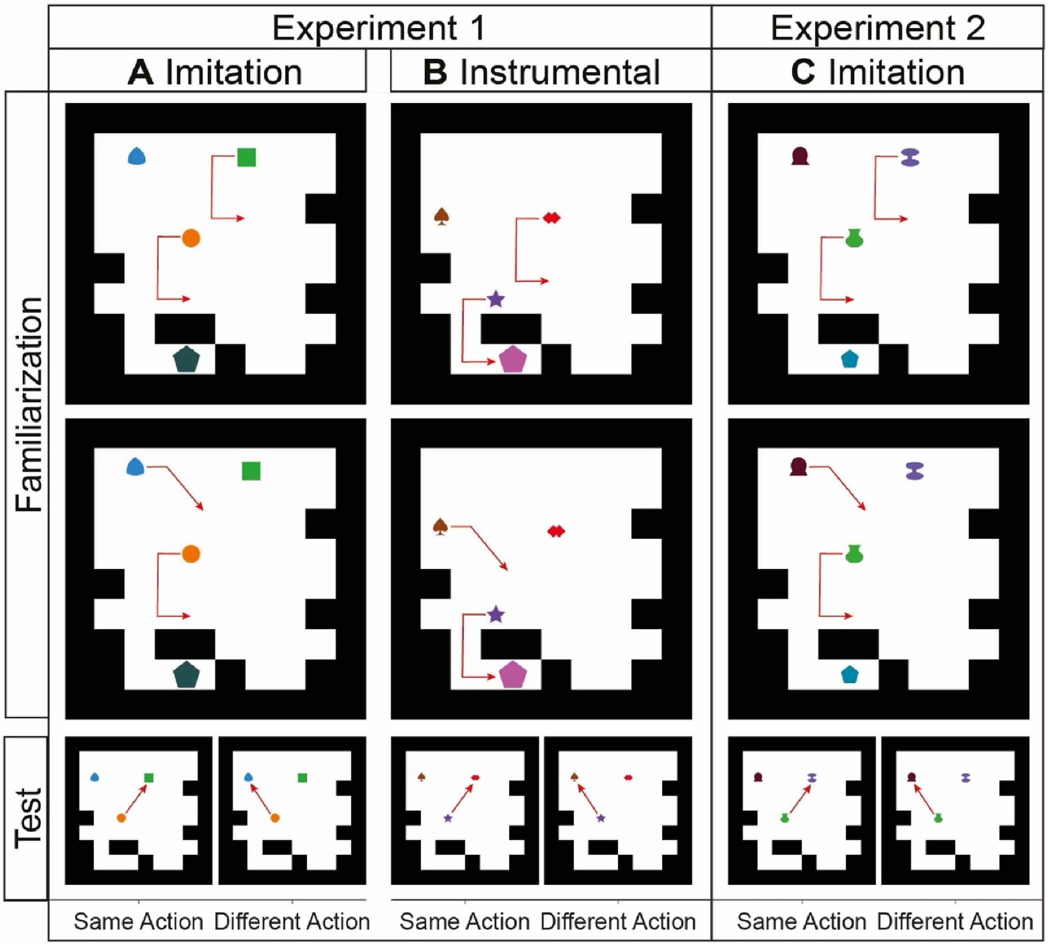

FIGURE 1 | Schematic of the displays used in Experiments 1 & 2. In the Imitation task of Experiment 1 (A), two kinds of familiarization trials were presented four times in alternation. In one kind of trial (top), the imitator (orange circle) performed the same action (e.g., a "C" shape) as one (e.g., a green square) of the two other agents, and in the other kind of trial (bottom), the third agent (e.g., a blue reuleaux triangle) performed a different action (e.g., a diagonal shape), after which the imitator continued to perform its original action. The agents' paths are shown in red for display purposes only. The Instrumental task of Experiment 1 (B) had the same structure and actions as the Imitation task, but the imitator's movement included efficient action around an obstacle in the environment to contact a stationary object (a teal pentagon), which changed color when it was touched. At test, the imitator approached either the agent it had performed the same action as or the agent it had performed a different action from. The Imitation task of Experiment 2 (C) was similar to that of Experiment 1, but, critically, included the agent changing color at the midpoint of its action to match the color change in the Instrumental task of Experiment 1.

then infants here would look equally long when the imitator approached either agent at test.

The Instrumental task otherwise matched the Imitation task except for minor, noncritical differences. The two tasks used different shapes (Figure 1), and those shapes had different absolute locations in the grid world during familiarization: While the grid-world layout and the relative locations of the three shapes were matched between the two tasks, the locations of the shapes were translated down in the grid world in the Instrumental task to allow the imitator to reach and contact the object. In both tasks and across both familiarization and test videos, the starting location of the imitator was always equidistant from the starting locations of the other two agents. The exact locations of the agents were matched across the two tasks in the test videos by averaging their locations from the first imitative familiarization video from the two tasks. In both tasks, the side of the agent whose action was imitated (left or right), the identity of that agent (Imitation: blue reuleaux triangle or green square; Instrumental: brown spade or red figure-eight), the agent that moved in the first familiarization video (the one whose action was or was not imitated), the action that was imitated (C-shape or diagonal), and the order of the two test videos (same-action first or different-action first) were counterbalanced.

2.1.3 | Procedure

Infants participated via scheduled Zoom sessions live with an experimenter. In the first 10 min of each session, the experimenter guided caretakers through setting up their computer, positioning their infant on their lap or in a highchair, and removing any distractions from their surroundings. The experimenter also instructed caretakers to close their eyes to not interfere with their infant's looking. The testing session was recorded through the Zoom recording function, capturing both the infant's face and the screen presenting the stimuli.

Each trial began with a 3 s attention grabber (a swirling blob in the center of the screen with a chiming sound). The experimenter, masked to the trial number, the order of the test trials, and what infants saw, coded infants' looking times live from the start of each video, and this live coding controlled the stimulus presentation through PyHab-online (Kominsky 2019) and slides.com. Each video played on loop until the infant looked away for 2 s consecutively (after looking for at least 1 s) or after a maximum of 60 s. A second coder, also masked to the trial number, the order of the test trials, and what the infants saw, recoded 8 randomly chosen sessions (25%) from each task, and the reliability between the first and second coders was high (ICC = 0.99).

3 | Results

We evaluated infants' performance with mixed-model linear regressions, and we obtained *p*-values using Type 3 Wald tests. The main results are shown in Figure 2A–B. We first evaluated infants' performance in each condition, including raw looking time as the dependent variable, outcome (approaching the agent it performed the same action as or approaching the agent it performed a different action from) as a fixed effect, and participant as a random-effects intercept. We found that infants looked significantly longer to the different-action (unexpected) outcome in the Imitation task ($F[1, 32] = 12.88$, $p = 0.001$), but we did not find a difference in their looking time to the two outcomes in the Instrumental task ($F[1, 31] = 0.14$, $p = 0.713$). We then compared the tasks directly, including outcome and task as fixed effects and participant as a random-effects intercept. We did not find a main effect of task ($F[1, 124] = 0.28$, $p = 0.597$) or outcome ($F[1, 83] = 2.18$, $p = 0.144$), but critically, we found a significant task X outcome interaction ($F[1, 83] = 4.47$, $p = 0.038$), with infants in the Imitation task, not the Instrumental task, looking longer to the different-action test outcome than the same-action test outcome. An analysis with the 17 infants who completed both tasks corroborated these results, with infants looking longer to the different-action test outcome in the Imitation task ($F[1, 16] = 5.06$, $p = 0.039$) but not in the Instrumental task ($F[1, 16] = 1.03$, $p = 0.326$).

4 | Discussion

Across two tasks presenting simple, minimally contrastive scenarios, we found that 15-month-olds understood an agent's imitative action as either social in one context or instrumental in another context. Moreover, these older infants attributed such social and instrumental goals to highly minimal agents: The agents in our study were simple shapes without eyes or limbs interacting in dyads via simple actions without sound in a grid world. Infants' differential performance in the Imitation and Instrumental tasks, moreover, suggests that infants understood the shapes as representing agents and objects and did not form their expectations based merely on similarities between shapes' actions. Infants' sensitivity to the social goals driving these agents' acts of imitation was consistent, moreover, with infants' sensitivities in prior studies using richer displays including: animations with more cues to agency and the imitation of actions and sounds (Powell and Spelke 2018a); recorded vignettes with puppets and people (Kudrnova, Spelke, and Thomas 2024); events in which animated agents imitated in the contexts of social groups (Powell and Spelke 2013); and live demonstrations (Agnetta and Rochat 2004; Meltzoff and Moore 1999).

While the Imitation and Instrumental tasks of Experiment 1 were closely matched, the null findings in the Instrumental task nevertheless raised the possibility that a difference between the tasks, besides the critical difference in intentional cues, might have driven the results. In particular, while the object in the Instrumental task changed color when the agent contacted it, there was no color change in the Imitation task. This color change in the Instrumental task may have distracted infants, preventing them from making any predictions about the agent's goals or subsequent actions. In Experiment 2, we aimed to replicate and extend the results of the Imitation task of Experiment 1 with an additional control to better match the attentional demands of Experiment 1's Instrumental task. Experiment 2 thus presented a new group of 15-month-old

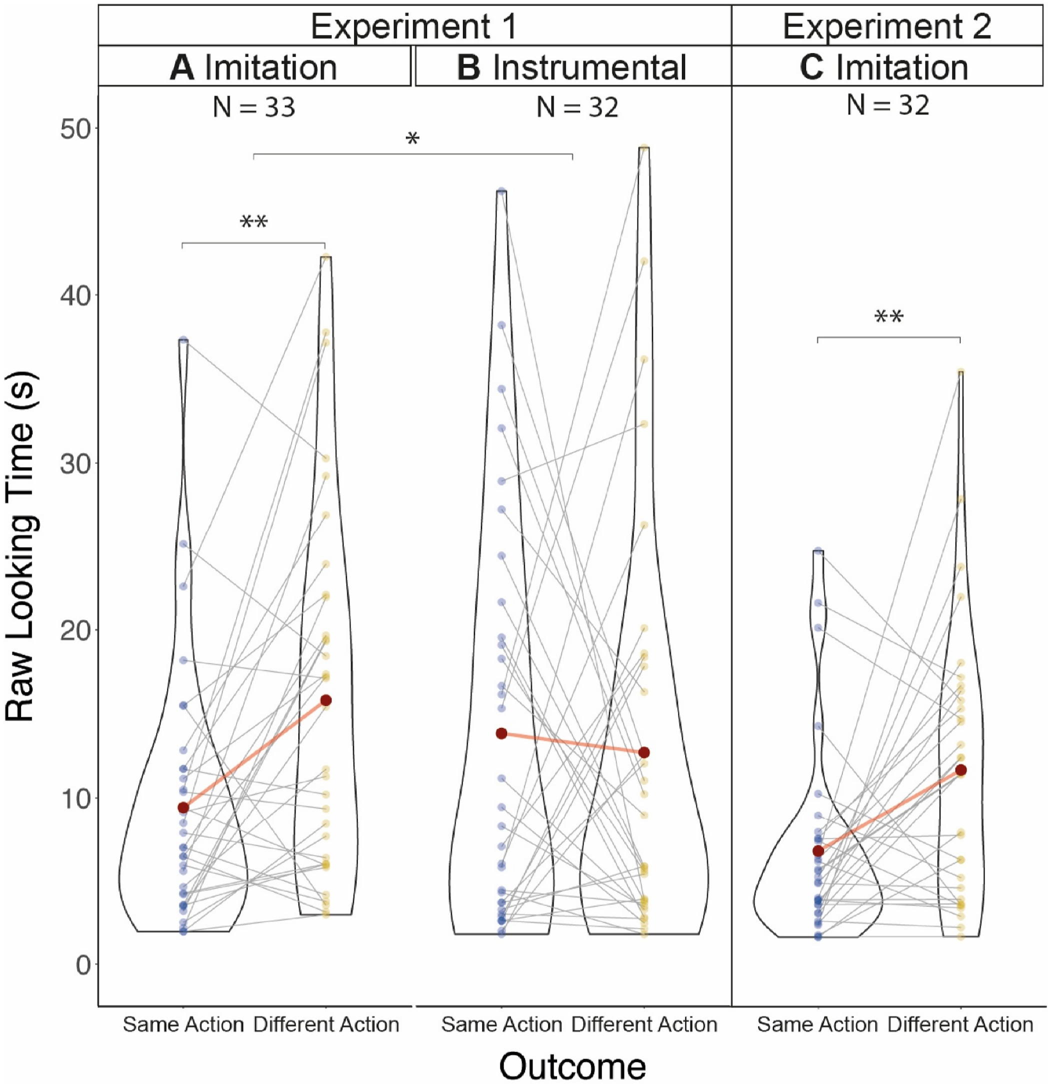

FIGURE 2 | Raw looking times to each outcome in the Imitation (A) and Instrumental (B) tasks of Experiment 1 and the Imitation (C) task of Experiment 2. Gray lines connect participants' looking times, represented by blue and yellow dots. Red dots and lines represent average looking times to each outcome for each task. $*p < 0.05$, $**p < 0.01$.

infants with a modified version of the Imitation task from Experiment 1, which included a color change to the imitator at the midpoint of its action.

5 | Experiment 2

5.1 | Methods

5.1.1 | Participants

Following Experiment 1, we planned a sample size of $N = 32$. Thirty-two typically developing 15-month-olds (*M*age: 15.01 m, range: 14.50 m–15.46 m, 19 girls) born at ≥ 37 weeks gestational age participated. An additional 29 sessions were excluded by a masked experimenter based on the same, predetermined criteria as Experiment 1, including: not completing the session (6); technical failure (6); poor video quality and/or missing video (2); caretaker interference (10), and extreme values (> 50 s) to one test trial, identified through Cook's Distance (5). Informed consent was obtained prior to each session, and the use of human participants for this study was approved by the Institutional Review Board at New York University.

5.2 | Materials & Procedure

The materials and procedure were the same for Experiment 2 as for the Imitation task of Experiment 1 except for the following changes. First, the imitator changed color at the midpoint of its action. We decided that the imitator, instead of the object, would exhibit the color change because: the object changing color without being contacted could have confused infants; the object changing color may have nevertheless looked causal given the temporal coincidence with the imitator's action; and the object was in a different part of the grid world than the imitator, and so changing the object's color could have drawn infants' attention away from the imitator's action. We chose to change the agent's color at the midpoint of its action because it matched the time point in the trial of the color change in the Instrumental task.

Additional minor changes to the stimuli included: using different shapes and a color-blind-friendly color palette for the agents (targets: purple dumbbell and brown crystal ball; imitator: green vase); a slightly smaller shape representing the object to match the size of the shapes representing the agents; making the actions 1.11*x* faster overall to maintain infants' attention throughout the entirety of the trials; and fixing the shape of the target agents' actions to only one side of the grid world while maintaining all other counterbalancing (the C-shape action always occurred on the right side of the grid world while the diagonal action always occurred on the left side of the grid world). While we found no interaction between outcome, action shape, and side of the grid world on infants' looking times in the Imitation task of Experiment 1 (mixed-model linear regression on raw looking time with outcome, action shape, and side as fixed effects and participant as a random-effects intercept: $F[1, 32] = 0.18$, $p = 0.673$), visual inspection of the task's videos suggested that fixing the side of each kind of action best prevented against the interpretation that a target might be approaching the imitator.

A second coder, also masked to the trial number, the order of the test trials, and what the infants saw recoded 8 randomly chosen sessions (25%) from each task, and the reliability between the first and second coders was high (ICC = 0.99).

6 | Results

We analyzed the data as in Experiment 1, and the results are shown in Figure 2C. In a mixed-model linear regression, we included raw looking time as the dependent variable, outcome (approaching the agent it performed the same action as or approaching the agent it performed a different action from) as a fixed effect, and participant as a random-effects intercept. As in the Imitation task of Experiment 1, we found that infants looked significantly longer to the different-action (unexpected) versus same-action (expected) outcome ($F[1, 31] = 8.54$, $p = 0.006$).

7 | General Discussion

In Experiment 1, we found that 15-month-olds understood an agent's imitative action as either social in one context or instrumental in another context when these contexts were simple and minimally contrastive. In Experiment 2, we replicated and extended the findings of Experiment 1 by showing that infants maintained their positive expectations about the social goal of an imitator's action even with the addition of a color change to the displays, matching the color change in the displays of the Instrumental task of Experiment 1. Together, these results suggest that from at least late infancy, we humans may have highly abstract notions of others as either potential social partners, whose actions have social goals, or as rational and efficient agents, whose actions are goal-directed toward objects.

On the one hand, infants' performance in these tasks may seem rather impressive, demonstrating differential and flexible interpretations of nearly identical scenarios. In particular, the agent's actions in the Instrumental task may have been open to multiple interpretations with varying degrees of likelihood (Jara-Ettinger et al. 2016; Hamlin et al. 2013; Powell 2022), including interpretations in which that agent's actions were nonsocial, social, or both: maybe the agent merely wanted to contact the object (nonsocial); maybe the agent merely wanted to imitate the other agent (social); maybe the agent wanted to move to the object *because* it involved making a social, imitative action (both); or maybe the agent wanted to imitate socially *and* reach the object efficiently and effectively (both). Nevertheless, infants favored what we adults might think was the most intuitive interpretation, especially given the multiple intentional cues (obstacles, efficiency, and causal effectiveness) in the Instrumental task's displays: that the agent's actions were indeed nonsocial (e.g., Schachner and Carey 2013). Future studies might explore whether and how infants evaluate agents' social versus instrumental goal attribution by varying the kind and number of intentional cues. For example, would infants recognize *inefficient*, imitative action toward a goal object as social? Previous research on older infants' and young children's understanding of ritualistic actions, as reviewed above, suggests yes and further indicates that cues like causal opacity, goal-demotedness (i.e., when there is no external motivation for acting on an object), and start- and end-state equivalence may more strongly signal a social versus object goal of an object-directed, imitative action (Legare et al. 2015; Liberman, Kinzler, and Woodward 2018).

On the other hand, infants' performance in these tasks is also open to a second interpretation, especially considering younger infants' limited understanding of the social value of object-directed imitation. As reviewed above, for example, older but not younger infants look longer and smile at an individual who has imitated their own actions on an object versus someone who has not (Agnetta and Rochat 2004; Carpenter, Uebel, and Tomasello 2013; Meltzoff 1988; Meltzoff and Moore 1999), and older infants imitate the actions of a social partner on an object (Meltzoff 1988). Infants' performance in the Instrumental task might thus be explained by a limit to (as opposed to flexibility with) human's early interpretations of apparently prosocial actions like imitation: When an object goal is present, infants may be unable to recognize any social value in an agent's imitation. The present tasks raise this possibility but cannot address it directly. For example, the null findings in the Instrumental task cannot specify whether infants attributed any goal to the agent at all, whether it be an object goal, a weakly social goal, or both, or whether infants evaluated the relative likelihoods of different social versus nonsocial goals for the agent. Future studies might explore what potential positive expectations infants can form about object-directed imitation by probing, for example, infants' expectations about an imitator's preference between the target agent of its imitation and the kind of object it had acted on.

Infants' potential failure to attribute any social goals to imitators who act on objects may even be analogous to their failures in other, seemingly unified early emerging domains of knowledge that are nevertheless composed of separate sets of early sensitivities, each with their own signature limits. For example, when infants are tasked with comparing quantities of 1 versus 4 sequentially presented items, they fail even though the items' ratio falls within their approximate numerical sensitivity because four exceeds the limits of their ability to track individual objects (Feigenson and Carey 2005). Studies of the perceptual and neural mechanisms supporting numerical cognition show that traces of such limits are present even in educated adults who have learned concepts of natural number (Hyde and Wood 2011; Intriligator and Cavanagh 2001). The 15-month-olds in the present experiment may thus have been subject to the same limits as young infants in understanding any social value of an imitator's action on an object, here persisting either into third-person judgments or for highly minimal social contexts. Indeed, traces of such limits in the social domain may even persist into adulthood: Our clearest social actions in adulthood are often those that involve no objects at all, like smiling, waving, or dancing (Schachner and Carey 2013; Spelke 2022). If our concepts of others in adulthood are composed of thinking of others' actions in these dissociable ways, either as socially driven or object-driven, then such limits might also underlie certain tensions in our treatment of others when the two ways of thinking about people compete for attention (Spelke 2022), resulting in moral dilemmas and other cruxes in ethical theories (Dillon 2024; Reilly 2019).

Nevertheless, in our eventual ability to think about others as both social and instrumental, we might still ask whether infants can see actions as simultaneously social and instrumental and what cues children and adults might use to simultaneously imbue social and instrumental meaning to others' actions. Future studies might thus explore what predictions infants make about agents, given myriad variations in both a target's and imitator's actions, goals, and efficiency. For example, do infants think two agents who have the same kind of goal object should affiliate, even if they make different unconstrained actions to reach these goals? Does an imitator contacting an object indicate a nonsocial goal more strongly than a target contacting an object? A systematic evaluation of a set of scenarios that probe candidate principles of early knowledge about social and instrumental actions and goals should be designed and evaluated as a suite or "benchmark," as has been done in a related area of infants' knowledge about the goals and rationality of agents when they are acting in goal-directed but nonsocial settings (Gandhi et al. 2021; Stojnić et al. 2023). The advantages of such an approach are numerous, including specifying how early principles of social intelligence might or might not cohere and permitting direct comparison between the performance of infants and computational models, a comparison which can uniquely specify the inductive biases at the foundations of social intelligence. The results of the present study and of such future explorations thereby promise to inform our understanding of the knowledge on which human social learning is built and thus the building of artificial intelligence that aims to interact with humans like humans do.

Author Contributions

Shannon Yasuda: conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, software, validation, visualization, writing – original draft, writing – review & editing. **Wenjie Li:** conceptualization, methodology, resources, software. **Deisy Martinez:** conceptualization, data curation, methodology, project administration. **Brenden M. Lake:** conceptualization, funding acquisition, methodology, writing – review & editing. **Moira R. Dillon:** conceptualization, funding acquisition, investigation, methodology, project administration, resources, supervision, validation, writing – original draft, writing – review & editing.

Acknowledgments

This work was supported by a National Science Foundation CAREER Award (DRL1845924 to MRD) and a DARPA grant on Machine Common Sense (HR001119S0005 to BML and MRD). We thank the generous families who volunteered their time to participate in this research.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Materials, methods, data, and analyses are available at: https://osf.io/qyfvm/.

References

Agnetta, B., and P. Rochat. 2004. "Imitative Games by 9-, 14-, and 18-Month-Old Infants." *Infancy* 6, no. 1: 1–36. https://doi.org/10.1207/s15327078in0601_1.

Bian, L., and R. Baillargeon. 2022. "When Are Similar Individuals a Group? Early Reasoning About Similarity and In-Group Support." *Psychological Science* 33, no. 5: 752–764. https://doi.org/10.1177/09567976211055185.

Buttelmann, D., N. Zmyj, M. Daum, and M. Carpenter. 2013. "Selective Imitation of In-Group over Out-Group Members in 14-Month-Old Infants." *Child Development* 84, no. 2: 422–428. https://doi.org/10.1111/j.1467-8624.2012.01860.x.

Carpenter, M., J. Uebel, and M. Tomasello. 2013. "Being Mimicked Increases Prosocial Behavior in 18-Month-Old Infants." *Child Development* 84, no. 5: 1511–1518. https://doi.org/10.1111/cdev.12083.

Csibra, G., G. Gergely, S. Bíró, O. Koós, and M. Brockbank. 1999. "Goal Attribution Without Agency Cues: The Perception of 'Pure Reason' in Infancy." *Cognition* 72, no. 3: 237–267. https://doi.org/10.1016/S0010-0277(99)00039-6.

Dillon, M. R. 2024. "Divisive Language." *Behavioral and Brain Sciences* 47: e124. https://doi.org/10.1017/S0140525X23003047.

Feigenson, L., and S. Carey. 2005. "On the Limits of Infants' Quantification of Small Object Arrays." *Cognition* 97, no. 3: 295–313. https://doi.org/10.1016/j.cognition.2004.09.010.

Gandhi, K., G. Stojnic, B. M. Lake, and M. R. Dillon. 2021. "Baby Intuitions Benchmark (BIB): Discerning the Goals, Preferences, and Actions of Others." *Advances in Neural Information Processing Systems* 34: 9963–9976.

Gergely, G., H. Bekkering, and I. Király. 2002. "Rational Imitation in Preverbal Infants." *Nature* 415, no. 6873: 755. https://doi.org/10.1038/415755a.

Gergely, G., and G. Csibra. 2020. "Sylvia's Recipe: The Role of Imitation and Pedagogy in the Transmission of Cultural Knowledge." In *Roots of Human Sociality*, edited by N. J. Enfield and S. C. Levinson, 229–255. London: Routledge.

Gergely, G., Z. Nádasdy, G. Csibra, and S. Bíró. 1995. "Taking the Intentional Stance at 12 Months of Age." *Cognition* 56, no. 2: 165–193. https://doi.org/10.1016/0010-0277(95)00661-H.

Green, P., and C. J. MacLeod. 2016. "SIMR: An R Package for Power Analysis of Generalized Linear Mixed Models by Simulation." *Methods in Ecology and Evolution* 7, no. 4: 493–498.

Hamlin, J. K., T. Ullman, J. Tenenbaum, N. Goodman, and C. Baker. 2013. "The Mentalistic Basis of Core Social Cognition: Experiments in Preverbal Infants and a Computational Model." *Developmental Science* 16, no. 2: 209–226. https://doi.org/10.1111/desc.12017.

Hyde, D. C., and J. N. Wood. 2011. "Spatial Attention Determines the Nature of Nonverbal Number Representation." *Journal of Cognitive Neuroscience* 23, no. 9: 2336–2351. https://doi.org/10.1162/jocn.2010.21581.

Intriligator, J., and P. Cavanagh. 2001. "The Spatial Resolution of Visual Attention." *Cognitive Psychology* 43, no. 3: 171–216. https://doi.org/10.1006/cogp.2001.0755.

Jara-Ettinger, J., H. Gweon, L. E. Schulz, and J. B. Tenenbaum. 2016. "The Naïve Utility Calculus: Computational Principles Underlying Commonsense Psychology." *Trends in Cognitive Sciences* 20, no. 8: 589–604. https://doi.org/10.1016/j.tics.2016.05.011.

Kominsky, J. F. 2019. "PyHab: Open-Source Real Time Infant Gaze Coding and Stimulus Presentation Software." *Infant Behavior and Development* 54: 114–119. https://doi.org/10.1016/j.infbeh.2018.11.006.

Kudrnova, V., E. S. Spelke, and A. J. Thomas. 2024. "Infants Infer Social Relationships Between Individuals Who Engage in Imitative Social Interactions." *Open Mind* 8: 202–216.

Legare, C. H., and M. Nielsen. 2015. "Imitation and Innovation: The Dual Engines of Cultural Learning." *Trends in Cognitive Sciences* 19, no. 11: 688–699. https://doi.org/10.1016/j.tics.2015.08.005.

Legare, C. H., N. J. Wen, P. A. Herrmann, and H. Whitehouse. 2015. "Imitative Flexibility and the Development of Cultural Learning." *Cognition* 142: 351–361. http://dx.doi.org/10.1016/j.cognition.2015.05.020.

Liberman, Z., K. D. Kinzler, and A. L. Woodward. 2018. "The Early Social Significance of Shared Ritual Actions." *Cognition* 171: 42–51. https://doi.org/10.1016/j.cognition.2017.10.018.

Liu, S., N. B. Brooks, and E. S. Spelke. 2019. "Origins of the Concepts Cause, Cost, and Goal in Prereaching Infants." *Proceedings of the National Academy of Sciences* 116, no. 36: 17747–17752. https://doi.org/10.1073/pnas.1904410116.

Meltzoff, A. N. 1988. "Infant Imitation After a 1-week Delay: Long-Term Memory for Novel Acts and Multiple Stimuli." *Developmental Psychology* 24, no. 4: 470–476. https://doi.org/10.1037/0012-1649.24.4.470.

Meltzoff, A. N., and M. K. Moore. 1999. "Persons and Representation: Why Infant Imitation Is Important for Theories of Human Development." In *Imitation in Infancy*, edited by J. Nadel and G. Butterworth, 9–35. New York: Cambridge University Press.

Muentener, P., and S. Carey. 2010. "Infants' Causal Representations of State Change Events." *Cognitive Psychology* 61, no. 2: 63–86. https://doi.org/10.1016/j.cogpsych.2010.02.001.

Powell, L. J. 2022. "Adopted Utility Calculus: Origins of a Concept of Social Affiliation." *Perspectives on Psychological Science* 17, no. 5: 1215–1233. https://doi.org/10.1177/17456916211048487.

Powell, L. J., and E. S. Spelke. 2013. "Preverbal Infants Expect Members of Social Groups to Act Alike." *Proceedings of the National Academy of Sciences* 110, no. 41. https://doi.org/10.1073/pnas.1304326110.

Powell, L. J., and E. S. Spelke. 2018a. "Human Infants' Understanding of Social Imitation: Inferences of Affiliation From Third Party Observations." *Cognition* 170: 31–48. https://doi.org/10.1016/j.cognition.2017.09.007.

Powell, L. J., and E. S. Spelke. 2018b. "Third-party Preferences for Imitators in Preverbal Infants." *Open Mind* 2, no. 2: 61–71. https://doi.org/10.1162/opmi_a_00018.

Reilly, B. J. 2019. *Getting the Blues: Vision and Cognition in the Middle Ages*. New York: Peter Lang.

Schachner, A., and S. Carey. 2013. "Reasoning About 'irrational' Actions: When Intentional Movements Cannot Be Explained, the Movements Themselves Are Seen as the Goal." *Cognition* 129, no. 2: 309–327. https://doi.org/10.1016/j.cognition.2013.07.006.

Skerry, A. E., S. E. Carey, and E. S. Spelke. 2013. "First-person Action Experience Reveals Sensitivity to Action Efficiency in Prereaching Infants." *Proceedings of the National Academy of Sciences* 110, no. 46: 18728–18733. https://doi.org/10.1073/pnas.1312322110.

Spelke, E. S. 2022. *What Babies Know: Core Knowledge and Composition*, Vol. 1. New York: Oxford University Press, Incorporated.

Stojnić, G., G. Kanishk, S. Yasuda, B. M. Lake, and M. R. Dillon. 2023. "Commonsense Psychology in Human Infants and Machines." *Cognition* 235: 105406. https://doi.org/10.1016/j.cognition.2023.105406.

Tomasello, M., M. Carpenter, J. Call, T. Behne, and H. Moll. 2005. "Understanding and Sharing Intentions: The Origins of Cultural Cognition." *Behavioral and Brain Sciences* 28, no. 5: 675–691. https://doi.org/10.1017/S0140525X05000129.

Woo, B. M., S. Liu, and E. S. Spelke. 2024. "Infants Rationally Infer the Goals of Other People's Reaches in the Absence of First-Person Experience With Reaching Actions." *Developmental Science* 27, no. 3: e13453. https://doi.org/10.1111/desc.13453.

Woodward, A. 1998. "Infants Selectively Encode the Goal Object of an Actor's Reach." *Cognition* 69, no. 1: 1–34. https://doi.org/10.1016/S0010-0277(98)00058-4.