Title: 15-month-olds' understanding of imitation in social and instrumental contexts

Authors: Shannon Yasuda[1]*, Wenjie Li[2], Deisy Martinez[1], Brenden M. Lake[1,2], Moira R. Dillon[1]*
*Correspondence may be addressed to sy3494@nyu.edu or moira.dillon@nyu.edu

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

Conflict of Interest Statement
The authors declare no conflicts of interest.

Data Availability Statement
Materials, methods, data, and analyses are available at: https://osf.io/qyfvm/?view_only=f31c779f2ff0447ca7f08651e76856e5

CRediT
All authors conceptualized the study, S. Y. and D. M. collected the data, S. Y. analyzed the data, S. Y. and M. R. D. wrote the original draft, B. M. L. provided edits, M. R. D. and B. M. L. secured funding, and M. R. D. supervised the study.

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

Word Count
2,994

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, external goals, and efficient, causally efficacious action. In showing that these older infants only imbue social value to imitation in the absence of such intentional cues, we suggest a signature limit to humans' early understanding of imitation. Our results suggest, 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.

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

Imitation of the expressions, goals, and actions of others is foundational to human social learning (e.g., Gergely & Csibra, 2020; Legare & 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 & Moore, 1999) and look longer at animated characters who imitate versus do not imitate other animated characters (Powell & Spelke, 2018b). Young infants also show consistent third-party predictions about the social value of others' novel but imitative sounds and actions (Powell & Spelke, 2018a) when judging both social group membership (Powell & Spelke, 2013) and dyadic relationships (Powell & 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, e.g., the character is "dancing." But when the character makes the very same series of actions to interact with objects, adults instead infer that the character's goal is object directed, e.g., the character is collecting and moving the objects (Schachner & Carey, 2013). How, then, might infants' social evaluation of imitative actions be affected by the presence of obstacles and external goals?

On the one hand, even young infants are keenly sensitive to obstacles, goals, and causal efficacy when interpreting other agents' 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 et al., 2013; Stojnić et al., 2023); that agents will show consistent object-based goals (Stojnić et al., 2023; Woo et al., 2024; Woodward, 1998); and that agents' actions on objects are causally efficacious (Liu et al., 2019; Muentener & 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 & Rochat, 2004; Carpenter et al., 2013; Meltzoff, 1988; Meltzoff & 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 et al., 2002; Meltzoff, 1988). Toddlers in the second year of life thus appear to become ever more attuned to the potential social value of simultaneously goal-directed and imitative 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 et al., 2018).

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 intentional cues such as obstacles, external goals, and efficient, causally efficacious 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. 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. In showing that these older infants only imbue social value to imitative actions in the absence of such intentional cues, we suggest a signature limit to humans' early understanding of the social value of imitation.

Methods

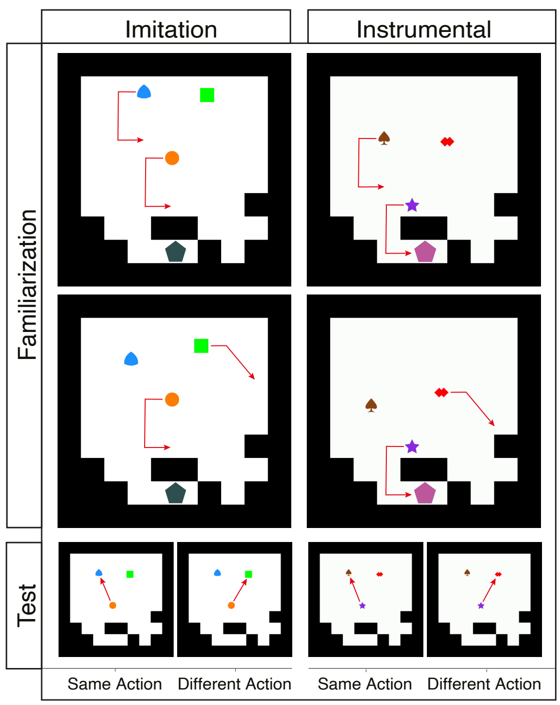

Figure 1. Schematic of the displays used in the Imitation and Instrumental tasks. In the Imitation task, two kinds of familiarization trials were presented four times in alternation. In one kind of trial (top), the imitator (orange circle) performed the same action (a "C" shape) as one (blue reuleaux triangle) of the two other agents, and in the other kind of trial (bottom), the third agent (blue square) performed a different action (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 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 (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.

Participants

Forty-seven typically developing 15-month-olds (Imitation task: N = 33, *M*age: 14.86m, range: 14.50m– 15.36m, 15 girls; Instrumental task: N = 32, *M*age: 14.93m, range: 14.53m–15.49m, 12 girls) born at ≥37 weeks gestational age participated in the study. Participants were asked to complete both an Imitation task and an Instrumental task during different Zoom sessions on different days within two 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. A total of N = 17 infants contributed data to both tasks. An additional thirty-four 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 < 1s to either test trial or < 2s 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 our university.

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 & 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 1**). 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 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 an agent it had imitated, then they would look longer to the different-action/unexpected test video.

The Instrumental task matched 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 efficacious (Liu et al., 2019), they could be seen as directed towards an object, not a social, goal (Schachner & Carey, 2013). If infants' predictions about the imitator's subsequent approach to one of the two agents were specific to contexts in which there are no obstacles or goal objects present, 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 shape whose action was or was not imitated), the action that was imitated (C-shape or diagonal), and the order of the test videos (same-action first or different-action first) were counterbalanced.

Procedure

Infants participated via scheduled Zoom sessions with a live experimenter. In the first ten minutes 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 infants' 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 3s 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 2s consecutively (after looking for at least 1s) or after a maximum of 60s. A second coder, also masked to the trial number, the order of the test trials, and what the infants saw on screen recoded 8 randomly chosen sessions (25%) from each task, and the reliability between the first and second coders was high (ICC = .99).

Results

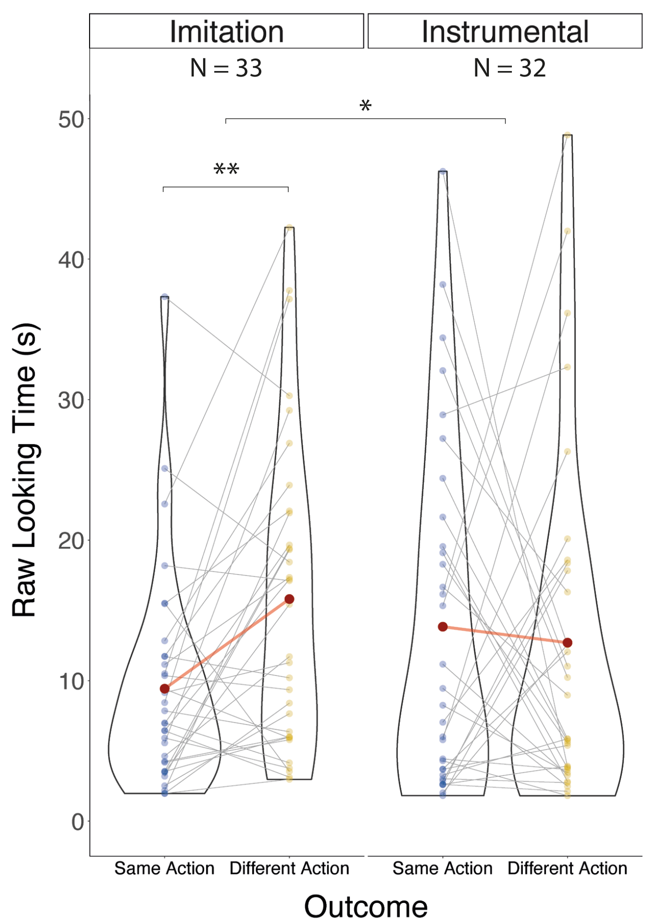

Figure 2: Raw looking times to each outcome in the Imitation and Instrumental tasks. Gray lines connect participants' looking times, represented by blue and yellow dots. Red dots and lines represent average looking times to each outcome for each task. * $p < .05$, ** $p < .01$.

We evaluated infants' performance on each task with mixed-model linear regressions, and we obtained *p*-values using Type 3 Wald tests. Our primary models 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. An additional model compared infants' looking times across the Imitation and Instrumental tasks, treating task as an additional fixed effect.

The main results are shown in **Figure 2**. We found that infants looked significantly longer to the different-action (unexpected) outcome in the Imitation task ($F(1, 32) = 12.88$, $p = .001$), but we did not find a difference in their looking time to the two outcomes in the Instrumental task ($F(1, 31) = 0.14$, $p = .713$). When comparing tasks directly, we did not find a main effect of task ($F(1, 124) = 0.28$, $p = .597$) or outcome ($F(1, 83) = 2.18$, $p = .144$), but critically, we found a significant task X outcome interaction ($F(1, 83) = 4.47$, $p = .038$), with infants in the Imitation task, not the Instrumental task, looking longer to the different-action test trial than the same-action test trial. 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 ($F(1, 16) = 5.06$, $p = .039$) but not in the Instrumental task ($F(1, 16) = 1.03$, $p = .326$).

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' performance was nevertheless consistent with the performance of infants in prior studies using richer displays including: animations with more cues to agency and the imitation of actions and sounds (Powell & Spelke, 2018a); recorded vignettes with puppets and people (Kudrnova et al., 2024); events in which animated agents imitated in the contexts of social groups (Powell & Spelke, 2013); and live demonstrations (Agnetta & Rochat, 2004; Meltzoff & Moore, 1999). 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 towards objects.

On the one hand, infants' performance in these tasks is impressive because of their flexible and different interpretations of nearly identical scenarios. But on the other hand, their performance may suggest a signature limit to human's early interpretations of apparently prosocial actions like imitation. In particular, the agent's actions in the Instrumental task were 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 socially imitate *and* reach the object (both).

The 15-month-olds in the present experiment may thus have been subject to the same limits as young infants in understanding a social, imitative action with an object goal, here persisting either into third-person judgments or for highly minimal social contexts. Indeed, traces of such limits may even persist into adulthood: Our clearest social actions even in adulthood may be those that involve no objects at all, like smiling, waving, or dancing (Schachner & 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, traces of these 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).

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. While the present study's Instrumental task was open to both social and nonsocial interpretations, several cues may have made its agent's actions seem decidedly less social. For example, there was only one, constrained path the agent could have taken to its goal object. Future studies might thus explore whether infants recognize unconstrained imitative movement to a goal object as relatively more social than constrained imitative movement to a goal object. Future studies might also explore what predictions infants might make about agents given other variations in the actions, goals, and efficiency of both the imitator's and target's actions. 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 their goals? 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 (Stojnić et al., 2023). The results of the present study and of such future explorations promise to inform our understanding of the foundational knowledge on which human social learning is built, as well as to aid the building of artificial intelligence that aims to learn like humans do.

References

Agnetta, B., & Rochat, P. (2004). Imitative Games by 9-, 14-, and 18-Month-Old Infants. *Infancy*, *6*(1), 1–36. https://doi.org/10.1207/s15327078in0601_1

Buttelmann, D., Zmyj, N., Daum, M., & Carpenter, M. (2013). Selective Imitation of In-Group Over Out-Group Members in 14-Month-Old Infants. *Child Development*, *84*(2), 422–428. https://doi.org/10.1111/j.1467-8624.2012.01860.x

Carpenter, M., Uebel, J., & Tomasello, M. (2013). Being Mimicked Increases Prosocial Behavior in 18-Month-Old Infants. *Child Development*, *84*(5), 1511–1518. https://doi.org/10.1111/cdev.12083

Csibra, G., Gergely, G., Bíró, S., Koós, O., & Brockbank, M. (1999). Goal attribution without agency cues: The perception of 'pure reason' in infancy. *Cognition*, *72*(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. doi:10.1017/S0140525X23003047

Gandhi, K., Stojnic, G., Lake, B. M., & Dillon, M. R. (2021). Baby Intuitions Benchmark (BIB): Discerning the goals, preferences, and actions of others. *Advances in Neural Information Processing Systems*, *34*, 9963–9976.

Gergely, G., Bekkering, H., & Király, I. (2002). Rational imitation in preverbal infants. *Nature*, *415*(6873), 755–755. https://doi.org/10.1038/415755a

Gergely, G., & Csibra, G. (2020). Sylvia's Recipe: The Role of Imitation and Pedagogy in the Transmission of Cultural Knowledge. In N. J. Enfield & S. C. Levinson (Eds.), *Roots of Human Sociality* (1st ed., pp. 229–255). Routledge. https://doi.org/10.4324/9781003135517-11

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*(2), 165–193. https://doi.org/10.1016/0010-0277(95)00661-H

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. https://doi.org/10.1016/j.tics.2016.05.011

Hamlin, J. K., Ullman, T., Tenenbaum, J., Goodman, N., & Baker, C. (2013). The mentalistic basis of core social cognition: Experiments in preverbal infants and a computational model. *Developmental Science*, *16*(2), 209–226. https://doi.org/10.1111/desc.12017

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

Kudrnova, V., Spelke, E. S., & Thomas, A. J. (2024). Infants Infer Social Relationships Between Individuals Who Engage in Imitative Social Interactions. *OPEN MIND*.

Legare, C. H., & Nielsen, M. (2015). Imitation and Innovation: The Dual Engines of Cultural Learning. *Trends in Cognitive Sciences*, *19*(11), 688–699. https://doi.org/10.1016/j.tics.2015.08.005

Liberman, Z., Kinzler, K. D., & Woodward, A. L. (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., Brooks, N. B., & Spelke, E. S. (2019). Origins of the concepts cause, cost, and goal in prereaching infants. *Proceedings of the National Academy of Sciences*, *116*(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*(4), 470–476. https://doi.org/10.1037/0012-1649.24.4.470

Meltzoff, A. N., & Moore, M. K. (1999). Persons and representation: Why infant imitation is important for theories of human development. In *Imitation in infancy.* (pp. 9–35). Cambridge University Press.

Muentener, P., & Carey, S. (2010). Infants' Causal Representations of State Change Events. *Cognitive Psychology*, *61*(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*(5), 1215–1233. https://doi.org/10.1177/17456916211048487

Powell, L. J., & Spelke, E. S. (2013). Preverbal infants expect members of social groups to act alike. *Proceedings of the National Academy of Sciences*, *110*(41). https://doi.org/10.1073/pnas.1304326110

Powell, L. J., & Spelke, E. S. (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., & Spelke, E. S. (2018b). Third-Party Preferences for Imitators in Preverbal Infants. *Open Mind*, *2*(2), 61–71. https://doi.org/10.1162/opmi_a_00018

Schachner, A., & Carey, S. (2013). Reasoning about 'irrational' actions: When intentional movements cannot be explained, the movements themselves are seen as the goal. *Cognition*, *129*(2), 309–327. https://doi.org/10.1016/j.cognition.2013.07.006

Skerry, A. E., Carey, S. E., & Spelke, E. S. (2013). First-person action experience reveals sensitivity to action efficiency in prereaching infants. *Proceedings of the National Academy of Sciences*, *110*(46), 18728–18733. https://doi.org/10.1073/pnas.1312322110

Spelke, E. S. (2022). *What Babies Know: Core Knowledge and Composition Volume 1*. Oxford University Press, Incorporated. http://ebookcentral.proquest.com/lib/nyulibrary-ebooks/detail.action?docID=7033700

Stojnić, G., Kanishk, G., Yasuda, S., Lake, B., M., & Dillon, M., R. (2023). Commonsense psychology in human infants and machines. *Cognition*, *235*, 105406. https://doi.org/10.1016/j.cognition.2023.105406

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, *28*(5), 675–691. https://doi.org/10.1017/S0140525X05000129

Woo, B. M., Liu, S., & Spelke, E. S. (2024). Infants rationally infer the goals of other people's reaches in the absence of first-person experience with reaching actions. *Developmental Science*, *27*(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*(1), 1–34. https://doi.org/10.1016/S0010-0277(98)00058-4