Title
Infants expect an agent to choose a goal that can be reached at a lower cost

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors
Schlingloff-Nemecz, Laura
Csibra, Gergely
Tatone, Denis
et al.

Publication Date
2024

Peer reviewed

Infants expect an agent to choose a goal that can be reached at a lower cost

Laura Schlingloff-Nemecz (laura.schlingloff@gmail.com)
Cognitive Development Center, Department of Cognitive Science, Central European University,
Vienna, Austria
TUM School of Social Sciences and Technology, Technical University of Munich,
Munich, Germany

Gergely Csibra (CsibraG@ceu.edu)
Cognitive Development Center, Department of Cognitive Science, Central European University,
Vienna, Austria
Department of Psychological Sciences, Birkbeck, University of London,
London, UK

Denis Tatone (denis.tatone@gmail.com)
Cognitive Development Center, Department of Cognitive Science, Central European University,
Vienna, Austria

Barbara Pomiechowska (b.pomiechowska@bham.ac.uk)
School of Psychology, University of Birmingham, Edgbaston, Birmingham, UK

Abstract

According to prominent accounts of early action understanding, infants' interpretation of others' actions is undergirded by an assumption of utility maximization. However, it is unclear whether this assumption applies only to selection among actions or also to selection among goals. Here, using an eye-tracking paradigm, we investigated whether 14- to 16-month-old infants would predict an agent to choose a lower-cost option when faced with two identical outcomes that could be reached at different costs. Infants directed more looks to the lower-cost option, and this effect was not merely due to visual saliency. These findings corroborate the proposal that infants rely on utility maximization when reasoning about an agent's likely goal and provide evidence of an early ability to represent and compare alternatives in the context of goal attribution.

Keywords: infant social cognition; action understanding; naive utility calculus; eye-tracking

Introduction

According to the naive utility calculus theory (Jara-Ettinger, Gweon, Schulz, & Tenenbaum, 2016), people spontaneously interpret others' actions as if guided by utility maximization. Developmental evidence suggests that this mechanism of action understanding may be available early in life, allowing infants to draw inferences about other agents' goals (Csibra, Bíró, Koós, & Gergely, 2003; Gergely, Nádasdy, Csibra, & Bíró, 1995), preferences (Liu, Ullman, Tenenbaum, & Spelke, 2017), and even individual competences (Pomiechowska & Csibra, 2022). Here, we investigate whether infants use the naive utility calculus to predict an agent's likely goal in a context in which multiple alternative goals yielding different utility are present.

Prior research has suggested that infants can ascribe a goal preference to agents upon observing their behavior. When infants are familiarized with an agent repeatedly approaching one of two available objects, they expect the agent to continue approaching this target when the locations of the two objects are switched at test (Woodward, 1998). In contrast, when infants are familiarized to an agent approaching the only object in the scene, they do not expect this object to be reached for when a second novel object appears at test (e.g., Choi et al., 2018; Luo & Baillargeon, 2005). The standard explanation for these results is that infants ascribe a preference to the agent when there are two options to choose from, but fail to do so when there is only one target object, as they remain naïve with regards to how much the agent values this option relative to a novel one (Youjung Choi & Luo, 2023). Under such an account, the selective reaching for one object over the other suggests that the reached-for object yields a relatively higher reward for the agent, so the agent should continue to approach it at test.

Other studies explicitly demonstrate that infants can construct preference rankings from observing patterns of goal approach. For instance, Mou and colleagues (Mou, Province, & Luo, 2014) showed that 16-month-olds can perform transitive inferences, such that if an agent approached A over B and B over C, they expect the agent to approach A over C. Similarly, Liu et al. (2017) found that infants employ utility reasoning to infer how much an agent values a particular goal relative to another based on how much effort he is willing to invest to reach it.

In these tasks, for infants to ascribe a preference, they need to (1) consider the available goal options as well as the means for reaching them, (2) calculate the expected utility for each of these options, and (3) identify which option yields the highest utility. The information about how rewarding each option is to the agent is judged by observing the agent's behavior. In other words, infants infer the agent's utility

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

function based, for instance, on which of two options is chosen in a given context or on the costs the agent is willing to incur to reach it.

Here we address whether infants can perform the inverse of this operation: If information about the relative utility of different goal options is given, will infants assume that the agent will choose the higher-utility option? Past studies showed that infants use path length as a proxy of action cost, which we build on here (Csibra, 2008; Csibra et al., 2003). In these studies, there was generally only one goal object available, which infants expect agents to approach in an efficient manner, i.e., by selecting the shortest available path.

In the present study we operationalized differences in utility as differences in effort (proxied by traversing paths of different length) required to approach one of two identical target objects. According to the naive utility calculus theory, goal-directed agents should (all else being equal) be expected to pursue the goal that can be reached at lower cost.

Schlingloff et al. (Schlingloff, Tatone, Pomiechowska, & Csibra, 2020) did not find support for this hypothesis in a looking-time experiment with 10-month-olds. They found that infants failed to differentiate between scenarios where an agent approached the lower-cost target (i.e., an object that could be reached by jumping over a lower wall relative to another) and scenarios where the agent approached the higher-cost target (i.e., an object that could be reached by jumping over a higher wall relative to another). Several reasons could have contributed to this null result. First, it may be that the ability to compare the utility of alternative options is not present at this age, which would contradict the standard explanation of the findings in the Woodward paradigm. Other related studies finding positive evidence for this capacity have tested infants in the second year of life (Liu et al., 2022; Scott & Baillargeon, 2013; Liu et al., 2017). For instance, a study by Scott and Baillargeon (2013) also addressed the question whether infants would expect an agent to choose a goal that could be reached with less effort. Here, effort was implemented as number of steps in an action sequence. Sixteen-month-olds, in a between-subject looking time design, looked longer when the agent selected the costlier option. A second potential reason for the null results obtained by Schlingloff et al. (2020) is that infants may have rationalized the agent's behavior when approaching the high-cost option by assuming that the higher effort was offset by commensurately higher rewards to be gained from that particular object (Liu et al., 2017).

The present experiment was devised to explicitly address these explanations. First, we selected a sample of older infants (14- to 16-month-olds). Second, we used an eye-tracking design, whereby, instead of measuring looking time to a completed event, we tested action prediction by recording which of two potential goal options infants would gaze towards when it was ambiguous which one the agent would approach.

In our experiment, infants were familiarized with a scenario where an agent approached a goal object located behind a door. Sometimes the door and the object were at a relatively small distance from the agent, and sometimes at a greater distance (on the far side of the screen). At test, both doors and objects were present (in the same locations as during familiarization). The agent started moving, but the video froze before it became clear which object he would approach. From this point, we measured where infants looked on the screen. We predicted that if they expected the agent to choose the object which would require less effort to reach (i.e. less costly to reach), they should look more at the closer object.

However, this result could also be accounted for by a simpler explanation. Namely, infants may have simply looked at the agent at the onset of the trial and fixated on the first relevant element of the display they encountered while exploring the scene from this visual anchor. To rule out this alternative, we also devised a control condition: here, at test, the two target objects were removed, while the doors remained in the same locations as in the other test event. Given that without a goal, there is no utility to be maximized in its pursuit, we predicted that if infants rely on the naive utility calculus to interpret the agent's actions, they should direct an equal amount of looks at the closer and the further door. On the other hand, if the simpler explanation is correct and visual attention and saliency drive infants' looking behavior in this task, they should behave similarly in the two conditions ("two-goals" and "no-goal") and look at the closest relevant element of the display (goal object or door) in each case.

The present experiment

This experiment was preregistered on the OSF [https://osf.io/cws8z].

Methods

Participants We preregistered a Bayesian stopping rule to determine sample size, that is, data collection was meant to be concluded once one of the following criteria was satisfied: (i) either we collect 48 valid data sets, or (ii) the Bayes Factor in the cross-condition comparison (see "Analyses") becomes equal to or greater than 10. The Bayesian statistical framework allows for sequential sampling (Visser et al., 2024). The Bayes Factor was first calculated after 16 valid samples, and after every 8 samples thereafter.

The final sample consisted of 48 14- to 16-month-old infants (27 male, mean age: 443 days). An additional 28 infants participated in the experiment but were excluded from the analysis for inattentiveness, i.e. failing to provide a sufficient amount of valid on-screen data (n = 22), parental interference (n = 1), and due to being erroneously tested although the stopping rule criteria had been met (n = 5). The study received full ethical approval from the University's ethics board.

Apparatus Infants were seated in their caregiver's lap in a darkened, soundproof room, approximately 60 cm away from the monitor. Their gaze was recorded using a Tobii Pro Spectrum Eye Tracker with an integrated 23.8-inch-diagonal monitor (resolution: 1920 x 1080; refresh rate: 60 Hz). A custom-made Python program building on PsychoPy 2021.1.3 (Peirce et al., 2019) was used for calibration, stimuli presentation, and gaze data collection. The stimuli were 3D animated videos created with Blender animation software (Stichting Blender Foundation, 2018).

Procedure and Stimuli Caregivers were instructed to hold the infants by their hips without impeding their ability to attend or disengage from the screen. Caregivers wore opaque glasses for the duration of the experiment. Before each trial, an attention-getting clip was shown.

Infants viewed two types of familiarization trials (short and long path goal approach) and two types of test trials ("two-goals": both goal objects and doors were present in the same locations as before; "no-goal": the goal objects were removed and only the doors were present). Familiarization and test trials were presented in alternation in a blocked design, containing both two-goals and no-goal test trials. Infants first watched four familiarization trials (short and long approach, two of each type in ABBA sequence), then a test trial (two-objects or no-object), then two familiarization trials (in AB sequence), a test trial (the other type than was shown before), and so on (see Figure 1). Infants saw a maximum of 20 trials including 6 test trials, 3 per type.

Familiarization. The familiarization videos showed a small blue agent approaching a goal object (red ball). Initially, the agent was facing forward and located on a small, circular, light blue plate on the left side of the screen, approximately in the center along the vertical axis. In front of the agent was a grey wall of a slightly lower height than the agent, spanning all the way across the screen horizontally. The wall contained a bright green segment – a door. The goal object was always located in front of the door.

The videos began by the red ball expanding and contracting twice (2 s). The agent then turned toward the ball and hopped in the air (2 s). Following this, the green door flashed yellow once and sank into the floor (3 s). Once the door had completely disappeared, leaving behind a green stripe on the ground (as a reminder where the door had been) and an opening in the wall, the agent started moving and, passing through the gap in the wall, approached the ball (5 s). Upon reaching the ball, the agent once again hopped in the air (1 s). Key events (ball and agent movement, door flashing and sinking) were accompanied by sound effects.

There were two types of familiarization videos: short-path approaches and long-path approaches. In the former, the door and ball were located at a relatively shorter distance to the agent, at 3 units on the left-right axis. In the latter, the door and ball were further away, at 7 units. Both video types had equal duration, which meant that in the long-path approaches, the agent moved slightly faster. At the end of the video, the final frame of the event was left on screen for 1 s before the onset of the next trial.

Test. Infants saw two types of test trial videos: two-goals and no-goal videos. Both contained a layout which was similar to that of the familiarization videos, except that there were now two doors in the wall, at the same locations as they had been in familiarization (at 3 and 7 units, respectively). Additionally, in the two-goals video, there was a ball behind each of the doors, whereas in the no-goal video, there was no ball.

The two-goals test video began similarly to the familiarization videos, in that the two red balls contracted and expanded. They did so sequentially (2 s each), and each ball's movement was met with a hop from the agent, who turned toward the ball which had just moved (2 s each) and finally turned to face the midpoint between the two balls. The green doors then blinked yellow, each once individually and subsequently both at the same time, and then simultaneously descended (5 s). At this point, the agent started moving, but only until he reached the edge of the blue plate on which he was located at the onset of the video (2 s).

The no-goal test video was identical, except for the fact that there were no balls behind the doors. The agent first turned to face the midpoint between the two doors, which then blinked and sank. As a result, this test video had a shorter duration than the two-objects video (9 s compared to 15 s).

At the end of a test video, the final frame of the event was left on screen for 10 s, which constituted the measurement period.

We counterbalanced the order of familiarization trials (short-path approach or long-path approach first), the order of test trials (two-goals or no-goals first), and the order in which the two balls in the two-goals test trial moved (closer ball or further ball first).

Stimuli can be accessed at https://osf.io/fvwga.

Data processing and analysis Of primary interest for our hypothesis were the gaze data collected during the test measurement period (i.e., a 10-second freeze frame at the end of the trial, after the agent stopped moving but before he approached a goal). The eye-tracker recorded sampled binocular gaze data every 16.67 ms. Gaze coordinates (x and y) of each sample were averaged across the eyes. Samples for which the x or y coordinate was missing, respectively, were removed. If there was data from only one eye available, these coordinates were used.

We defined three areas of interest (AOI) on the screen: the agent (A), the closer ball or door (C), and the further ball or door (F). The main dependent variable was the proportion of looking at the closer ball or door (i.e., the total looking to the closer option divided by the sum of total looking to both options: propC = totalC / (totalC + totalF). We calculated an average value of propC per condition for each participant.

We performed two types of analysis on this variable. First, we directly compared the looking patterns in the two test trial types: to test whether infants on average looked longer at the closer option in the two-objects compared to the no-object trials, we used a one-sided Bayesian paired samples t-test, predicting that the mean propC would be higher in the former than the latter. Second, we compared the mean propC in each condition to chance: to test whether infants would expect the agent to approach the closer rather than the further option in both test trial types (or, as we hypothesized, only in the two-objects trials), we used Bayesian one-sample t-tests (one-sided). Since there is little prior knowledge about the current effects, we used a default prior option, a Cauchy distribution (scale = 0.707). Statistical analyses were performed in R version 4.3.2 (R Core Team, 2023).

To be included in the final sample, participants had to provide a valid data set, which minimally consisted of four valid familiarization trials and two valid test trials, one test trial of each type (two-objects and no-object). A valid familiarization trial was defined as a trial during which the participant attended to the goal-approach action at least until the agent had crossed the line on the ground that demarcated the opening in the wall. A valid test trial was defined as a trial in which the participant (1) contributed at least 50% of on-screen data during the initial, animated phase of the video, (2) contributed at least four cumulative seconds of on-screen data during the test measurement period, and (3) gazed for a minimum of 300 ms (i.e., congruent with fixation) to at least one of the balls/doors during the test measurement phase.

Further exclusion criteria were caregiver interference (e.g., talking to the infant, pointing at the screen), experimenter error, external distractors (e.g., outside noise), and technical failure.

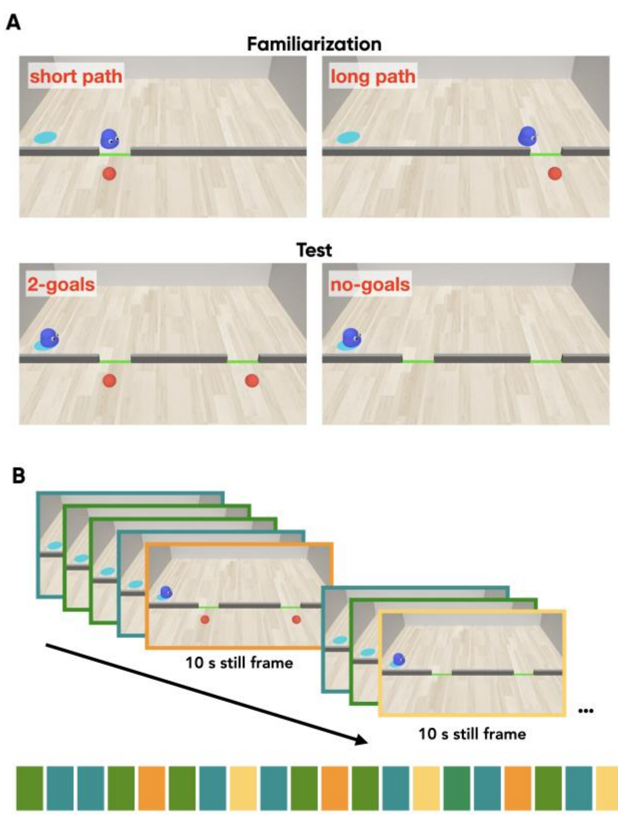

Figure 1: (A) Stimuli used in familiarization and test. During familiarization, the agent approached an object that was closer ("short path") or further ("long path") from his initial position. At test, the video froze on a still frame depicting a similar scene as during familiarization, now with both goal objects present ("two-goals") or none of them ("no-goal"). (B) Schematic depiction of testing procedure. Familiarization and test trials (max. 20 in total) were presented in blocks: Sets of familiarization trials interspersed with test trials, in which infants' gaze was recorded during a 10 s still frame at the end of the video. Familiarization trials are represented by green rectangles; test trials by orange rectangles.

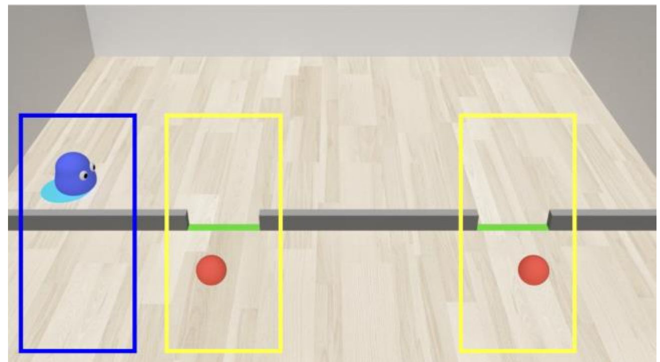

Figure 2: Screenshot from a two-goals test trial measurement phase. Rectangles represent the AOIs (yellow: the close (C) and far option (F), blue: the agent (A)).

Results

Hypothesis-driven results The mean proportion of looking to the closer option in the two-goals condition was mean(propC_two-goals) = .606 (SD = .204), and in the no-goal condition mean(propC_no-goal) = .535 (SD = .265). A direct comparison of the proportion of infants' looking to the closer option (propC) in the two-goals and the no-goal conditions showed no evidence that the looking behavior in the two events differed (BF: 0.8). However, when comparing propC to chance for each of the two conditions, we found that as predicted, infants looked longer at the closer option in the two-goals condition (BF: 73.6) but did not do so in the no-goal condition (BF: 0.4). Thus, one of the two preregistered analyses supported our hypothesis.

When conducting the analysis with the additional 5 infants who had been tested after the stopping criterion was already met and who were therefore excluded, results were similar. In the direct comparison, there was no evidence for a difference in looking patterns between conditions (BF: 0.8), but when comparing propC to chance, infants looked longer at the closer option in the two-goals condition (BF: 92.6) but not in the no-goal condition (BF: 0.4). With this inclusion criterion, the mean proportion of looking to the closer option in the two-goals condition was mean(propC_two-goals) = .6 (SD = .199) and in the no-goal condition mean(propC_no-goal) = .534 (SD = .253).

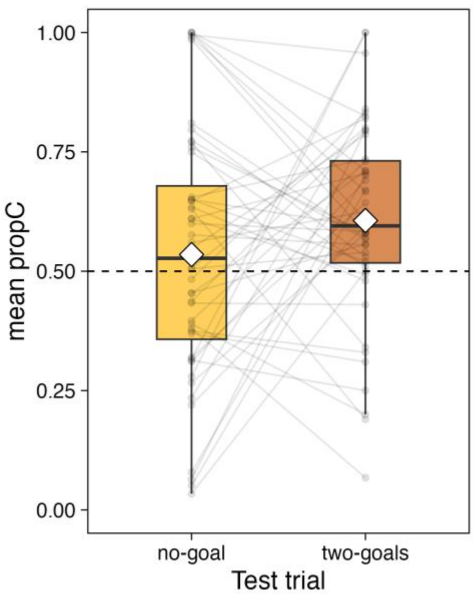

Figure 3: Box plot of mean proportions of looking to the closer option (propC) in no-goal and two-goals test trials. Light grey lines connect average looking proportions of individual participants, diamonds indicate means, horizontal lines indicate medians, boxes indicate middle quartiles, and whiskers indicate points within 1.5 times the interquartile range from the upper and lower edges of middle quartiles. The dashed horizontal line indicates chance (i.e., equal looking to the closer and further option).

Exploratory results During the test measurement period, infants attended to the screen for an average of 6.81 s (SD = 2.51 s) in the two-goals condition, and 6.68 s in the no-goal condition (SD = 2.25 s).

We assessed whether the proportion of looking at the agent differed in the two conditions and found that infants looked more at the agent in no-goal compared to two-goals condition (BF: 1231.6; two-sided test), which is not surprising given that the former contained fewer elements to look at.

We further looked at the proportion of looking to the closer option during the movie phase of the test trials. Here, infants looked more at the closer option in both two-goals (BF: 9152.8) and the no-goal condition (BF: 2480), and there was no evidence that this pattern differed between the conditions (BF: 0.3; all tests two-sided).

Discussion

The aim of the present experiment was to test whether 14- to 16-month-old infants would expect an agent to maximize his utility by choosing one of two identical goal options that could be reached at lower cost. To test this hypothesis, we familiarized infants with an agent approaching a goal object either behind a closer or a further door. Consistent with our main hypothesis, we found that, when the agent was presented with both objects at test, infants looked more at the closer object, suggesting that they predicted the agent approach it over the other. Corroborating the claim that infants did not fixate on the closer object simply because it represented the first salient item they encountered when scanning the scene from an agent-centered perspective, infants did not look longer at the wall closer to the agent when neither goal object was present at test.

We found support for this prediction in one of our two preregistered analyses: while there was no evidence that the looking patterns in the two conditions were different from each other when comparing them directly, infants looked to the closer option longer than expected by chance only in the two-goals condition when the goal objects were present. When the objects were absent, infants looked equally to the two doors regardless of their distance from the agent[1]. This result indicates that infants considered the closer ball, which was less costly to approach, a better candidate for a prospective goal object in a context where multiple options became available, and thus ascribed choice of a better option to the agent.

These results differ from those of Schlingloff et al. (2020), which provided no evidence that infants expected an agent to choose a lower-cost over a higher-cost goal. We hypothesized that two factors may have contributed to the different results: (1) In the present experiment, we tested older participants, as we reasoned that infants in the first year of life may not yet be able to possess the ability to compare alternative options, which is presupposed in the current task (even though this would be inconsistent with the widely held interpretation of infants' behavior in the Woodward paradigm). Compatibly with this claim, several studies providing explicit evidence for such a capacity were carried out with infants in the second year of life (Cesana-Arlotti, Varga, & Téglás, 2022; Duh & Wang, 2019; Liu et al., 2022; Mou et al., 2014; Scott & Baillargeon, 2013; Liu et al., 2017; Robson, Lee, Kuhlmeier, & Rutherford, 2014). (2) Instead of measuring looking times, we recorded the proportion of looks that infants directed to the two potential goal objects. Reasoning postdictively about whether an agent acted efficiently requires comparing the agent's observed behavior at test with the non-chosen alternative. Reasoning predictively about an agent's behavior, on the other hand, only requires future-directed hypothetical reasoning about which of two possible but unrealized outcomes the agent may be likely to bring about (Gerstenberg, 2022). Developmental research has shown that counterfactual reasoning may be more difficult for children than hypothetical reasoning (Kominsky et al., 2021; Robinson & Beck, 2000).

While the present results provide initial evidence for utility-based reasoning about the hypothetical cost of different possible actions, it should be noted that the control condition only ruled out the possibility that infants' gaze was primarily guided by any perceptually salient item closer to

[1] Note that Thomas et al. (2022), who also calculated infants' proportion of looking to one of two potential targets an agent may approach, only performed comparisons to chance in the experimental and control conditions, respectively.

the agent (such as the close door). There is, however, a second alternative that the control condition cannot account for: Infants may have scanned the scene searching specifically for target goal objects in the vicinity of the agent, and used physical distance as a heuristic to select a preferred option, rather than actually simulating and comparing the actions that would be required to bring about each goal. In other words, infants may have simply expected the agent to approach any goal object, and stopped looking once their gaze landed on a suitable candidate, which, due to its physical proximity, was more likely to be the closer ball. This account predicts that infants should look more to the closer of multiple goal objects, but produce a similar proportion of looks to the two doors when the doors are absent, because their search strategy is focused only on goal objects. Our results are compatible with both accounts, so future research will have to disambiguate between them, for instance, by conducting a follow-up experiment where the object that is closest to the agent as the crow flies is not the one that is the least costly to reach.

Regardless, our study provides tentative support for the proposal that infants in the second year indeed apply naive utility calculus to predict which potential goal an agent will choose. Beyond investigating the mechanisms of early action understanding, this study also contributes to ongoing debates on the developmental origins of modal reasoning (Leahy & Carey, 2020; Redshaw & Ganea, 2022). Our study is the first to use eye-tracking to demonstrate that infants expect agents to minimize their action cost (cf. Paulus et al., 2011), and points to promising directions for further research with this method.

Acknowledgments

We thank the families who participated in this experiment, and Dorottya Mészégető for help with data collection. This research has received partial funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 742231 (PARTNERS).

References

Cesana-Arlotti, N., Varga, B., & Téglás, E. (2022). The pupillometry of the possible: An investigation of infants' representation of alternative possibilities. *Philosophical Transactions of the Royal Society B.*, 377, 20210343.

Choi, Youjung, & Luo, Y. (2023). Understanding preferences in infancy. *WIREs Cognitive Science*, *14*, e1643.

Choi, You-jung, Mou, Y., & Luo, Y. (2018). How do 3-month-old infants attribute preferences to a human agent? *Journal of Experimental Child Psychology*, *172*, 96–106.

Csibra, G. (2008). Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition*, *107*, 705–717.

Csibra, G., Bíró, S., Koós, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science*, *27*, 111–133.

Duh, S., & Wang, S. (2019). Infants Detect Patterns of Choices Despite Counter Evidence, but Timing of Inconsistency Matters. *Journal of Cognition and Development*, *20*, 96–106.

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*, 165–193.

Gerstenberg, T. (2022). What would have happened? Counterfactuals, hypotheticals and causal judgements. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *377*, 20210339.

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*, 589–604.

Kominsky, J. F., Gerstenberg, T., Pelz, M., Sheskin, M., Singmann, H., Schulz, L., & Keil, F. C. (2021). The trajectory of counterfactual simulation in development. *Developmental Psychology*, *57*, 253–268.

Liu, S., Pepe, B., Ganesh Kumar, M., Ullman, T. D., Tenenbaum, J. B., & Spelke, E. S. (2022). Dangerous Ground: One-Year-Old Infants are Sensitive to Peril in Other Agents' Action Plans. *Open Mind*, *6*, 211–231.

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*, 1038–1041.

Luo, Y., & Baillargeon, R. (2005). Can a Self-Propelled Box Have a Goal?: Psychological Reasoning in 5-Month-Old Infants. *Psychological Science*, *16*.

Mou, Y., Province, J. M., & Luo, Y. (2014). Can infants make transitive inferences? *Cognitive Psychology*, *68*, 98–112.

Paulus, M., Hunnius, S., Van Wijngaarden, C., Vrins, S., Van Rooij, I., & Bekkering, H. (2011). The role of frequency information and teleological reasoning in infants' and adults' action prediction. *Developmental Psychology*, *47*, 976–983.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., … Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*, 195–203.

Pomiechowska, B., & Csibra, G. (2022). Infants infer motor competence from differences in agent-specific relative action costs. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *44*.

R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Robinson, E. J., & Beck, S. R. (2000). What is difficult about counterfactual reasoning? In *Children's Reasoning and the Mind*. Psychology Press.

Robson, S. J., Lee, V., Kuhlmeier, V. A., & Rutherford, M. D. (2014). Infants use contextual contingency to guide their interpretation of others' goal-directed behavior. *Cognitive Development*, *31*, 69–78.

Schlingloff, L., Tatone, D., Pomiechowska, B., & Csibra, G. (2020). Do infants think that agents choose what's best? *Proceedings of the Annual Meeting of the Cognitive Science Society*, *42*, 1495–1501. Cognitive Science Society.

Scott, R. M., & Baillargeon, R. (2013). Do Infants Really Expect Agents to Act Efficiently? A Critical Test of the Rationality Principle. *Psychological Science*, *24*, 466–474.

Visser, I., Kucharský, Š., Levelt, C., Stefan, A. M., Wagenmakers, E.-J., & Oakes, L. (2024). Bayesian sample size planning for developmental studies. *Infant and Child Development*, *33*(1), e2412.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*, 1–34.