Infants do not use payoff information to infer individual goals in joint-action events

Tatone, Denis; Schlingloff-Nemecz, Laura; Pomiechowska, Barbara

DOI:
10.1016/j.cogdev.2023.101329

Cognitive Development 66 (2023) 101329

Contents lists available at ScienceDirect

Cognitive Development

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

Infants do not use payoff information to infer individual goals in joint-action events

Denis Tatone *, Laura Schlingloff-Nemecz, Barbara Pomiechowska

Department of Cognitive Science, Central European University, Vienna, Austria

ARTICLE INFO

Keywords:
Infancy
Goal attribution
Joint action
Helping
Reward distribution

ABSTRACT

For observers the occurrence of a joint action (JA) is inherently ambiguous with respect to the goals of the individuals involved. People may work together, for instance, because they are motivated to obtain material rewards or to help others. We hypothesized that to solve this interpretive ambiguity observers leverage information about the JA's payoff structure. Specifically, when a JA yields material rewards for a participating agent (as well as their partner), their behavior can be straightforwardly explained as instrumental to the obtainment of these rewards. Conversely, when a JA does not yield material rewards for the agent (but does so for her partner), the unrewarded agent's contribution needs to be accounted for by positing other types of goals, such as assisting her partner in obtaining her rewards. We examined this hypothesis across three looking-time experiments with 12-month-olds: specifically, we tested whether the absence of material rewards for an agent participating in a JA would prompt infants to interpret her participation as prosocially motivated. Consistent with this hypothesis, Experiment 1 showed that, after having been familiarized to two dyadic JA events resulting in one or both agents being rewarded, infants selectively expected the unrewarded agent to act altruistically towards her former JA partner by giving her a resource. Experiments 2–3 examined whether this expectation was driven by the prosocial interpretation of the unrewarded agent's behavior or by changes in the number or distribution of resources between familiarization and test. Contradicting the hypothesis that infants interpreted the agent's behavior as prosocially motivated, in Experiment 2 we observed similar looking times when the unrewarded agent performed a prosocial action (giving) or an antisocial action (taking) towards her partner at test. Further, in a close replication of the original experiment in which the change in the number of familiarized objects occurred in the test event featuring the rewarded agent rather than the unrewarded one (Experiment 3), infants produced a looking-time pattern opposite to the one first obtained. Taken together, these findings suggest that infants encoded the payoff structure of JA events (i.e., the number/distribution of resources that the interaction brought about) but did not leverage this information to infer the individual goals of participating agents. The present evidence calls for a critical re-evaluation of our original hypothesis and for further research into the mechanisms by which infants disambiguate the motives of agents involved in joint actions.

* Corresponding author.
E-mail address: TatoneD@ceu.edu (D. Tatone).

https://doi.org/10.1016/j.cogdev.2023.101329
Received 2 July 2022; Received in revised form 26 February 2023; Accepted 22 March 2023
Available online 5 April 2023

1. Introduction

Coordination is ubiquitous in our daily lives, yet the reasons why people decide to work together towards a certain goal are often varied. Some instances of coordination emerge as a byproduct of overlapping individual goals (e.g., people queuing in front of a cashier). Others entail working together towards a common reward (e.g., two friends picking mushrooms to be shared over dinner). Others still are supported by prosocial intentions (e.g., a person working with her neighbor to help her repaint the fence). As these examples illustrate, coordination cues, such as acting similarly or synchronously on the same target, span across different types of interactions, whether selfish, mutualistic, or altruistic. This creates an interpretive quandary: if coordination cues are themselves insufficient to indicate specific interaction types, how can naïve observers reliably infer the reasons why others decided to participate in a joint endeavor?

One way to solve this problem is by leveraging information about the payoff structure of a joint action (henceforth, JA): i.e., the distribution of rewards that a JA brings about. Imagine two siblings, Deco and Malvin, pulling a heavy chest out of a basement. Watching this scene, it would be reasonable to infer that these two agents are willing to act together to realize a certain outcome (e.g., getting the chest out), but this evidence itself is insufficient to determine the individual motives that prompted either sibling to join forces and act together. Perhaps Deco hoped to find one of his childhood toys inside the chest, or maybe he sought to help Malvin retrieve one of hers. Knowing who benefitted from the chest's contents provides us with useful evidence to arbitrate among these interpretations. Namely, when a JA produces material rewards for an agent, her involvement can be justified simply by appealing to the obtainment of these rewards (e.g., Malvin pulled the chest to get her toy). In contrast, when a JA does not bear material rewards for said agent but does so for her partner, other rewards need to be posited to justify the agent's participation efforts. Given the interactive context in which a JA takes place, a plausible interpretation of the unrewarded agent's behavior is to help her partner fulfill her own goal (e.g., Malvin pulled the chest to help Deco get her toy). In other words, while the presence of material rewards may induce observers to frame an agent's participation in a JA as directed to a *self-interested* goal (i.e., to acquire material rewards), their absence should encourage an interpretation of the agent's behavior as directed to a *prosocial* goal (i.e., to help the partner acquire these rewards). Importantly, depending on the interpretation adopted, the observation of the JA licenses critically different inferences about an agent's broader disposition. For instance, if Malvin is helping Deco, we may be justified in inferring that she may be willing to continue benefiting his brother in other situations. In contrast, no such inferences would be warranted if Malvin's participation in the JA can be explained as directed to acquiring material rewards.

Building on these premises, we examined whether the presence of material rewards modulates the interpretation of an agent's participation in a JA. We targeted 12-month-olds because they have the capacity to apprehend an agent's behavior as directed to either self-interested or prosocial ends, as evidenced by their understanding of actions such as acquiring resources (taking: Eason et al., 2018), providing them to others (giving: Tatone et al., 2015), or assisting others in fulfilling their goals (helping: Hamlin et al., 2007). Moreover, by 9 months of age, infants can infer the shared goal of simple collaborative interactions consisting of similar and simultaneous actions (e.g., pushing an object in unison towards a certain location: Begus et al., 2020). This initial competence eventually expands to encompass more complex joint endeavors involving multi-step and complementary actions by around 18 months of age (e.g., one agent opens a box, while another retrieves an object inside: Henderson & Woodward, 2011).

Although the issue of how to arbitrate among multiple goal hypotheses is ubiquitous to action interpretation at large (for discussions in the domain of instrumental and social actions, see: Csibra & Gergely, 2007; Tatone et al., 2015), an explicit treatment of this problem has been lacking in the developmental literature on JA. The research on early understanding of JA has primarily focused on whether infants could bind two agents' actions in a collaborative sequence directed towards one distal goal (e.g., placing an object in a certain location: Fawcett & Gredebäck, 2013, 2015; Mascaro & Csibra, 2022). This evidence shows that infants can interpret individual contributions as directed to a shared outcome based on specific properties of their patterning (e.g., whether they are spatiotemporally contingent or jointly efficient in bringing about the outcome) but falls short of elucidating whether and how they disambiguate among the possible individual goals supporting JA. As we argued, the relation between individual goals and the production of a shared outcome is overdetermined (i.e., many-to-one): in our example, Malvin may have joined Deco in pulling the chest for self-interested or prosocial ends alike. This ambiguity looms also in the researchers' framing of JA scenarios: for instance, superficially similar interactions involving two agents acting together on a box to retrieve an object inside have sometimes been interpreted as an example of altruism (e.g., helping, a second-order prosocial goal, Hamlin & Wynn, 2011; Hamlin et al., 2007) and other times as an example of mutualism (a joint first-order goal, Henderson & Woodward, 2011), despite the different expectations about reward distribution that these interpretations engender (in helping, rewards should accrue to the Helpee; in mutualism, to both agents).

Only a couple of studies to date examined the relation between JA and individual payoffs by exploring infants' expectations about benefit distribution following collaborative actions. In one such study, 14-month-olds exposed to the aforementioned box scenario expected the agent acquiring the object to share it with their partner instead of monopolizing it (Wang & Henderson, 2018a; for similar results using a block-stacking and a joint-chasing scenario, see: Wang & Henderson, 2018b; Vorobyova, 2021). These findings suggest that infants spontaneously use information about the participation costs of JA participants to generate inferences about their prospective benefits according to retributive principles (by which contributions should be matched with commensurate rewards). The current work, in contrast, investigates whether infants leverage information about reward distribution (specifically, whether an agent gained materially from a JA or not) to disambiguate the goals (self-interested or altruistic) supporting the agent's involvement in a JA.

To this end, we presented 12-month-olds with videos featuring two pairs of agents. Each pair pulled a tray baited with a different number of cookies (2 vs. 1) in a synchronous manner, but for one pair this interaction resulted in an equal payoff (1–1, i.e., one cookie for each agent) and for the other one it resulted in an unequal payoff (1–0, i.e., one cookie for only one agent). By manipulating the payoff structure of the JA without changing the dynamics of the interaction observed, we could isolate the effects of material rewards on action interpretation. To assess whether the absence of material rewards prompted infants to adopt a prosocial interpretation of an agent's participation in the JA, we tested whether infants specifically expected the unrewarded agent to act altruistically with her former JA partner by giving her an object. The rationale for adopting this test measure was as follows: if infants interpreted the unrewarded agent's participation in the JA as indicative of a prosocial disposition towards her partner, they should find a novel prosocial behavior (giving) targeting the same partner at test congruent with this construal.

Two methodological points should be highlighted here. First, we used JA scenarios consisting of synchronous parallel actions to facilitate the interpretation of the two agents as coordinating. This choice was motivated by the fact that young infants fail to infer the goals of JAs involving hierarchical and complementary roles (e.g., Henderson & Woodward, 2013), but succeed when presented with JAs featuring synchronous and similar actions (Begus et al., 2020). Second, we used spatially separated rewards to mitigate the expectation that the payoff of the JA itself was itself joint, and thus accruing to both participants (cf. Wang & Henderson, 2018a). Doing so allowed us to examine whether we could induce infants to shift their interpretation of an agent's participation in the JA (from self-interested to prosocial) by manipulating their individual payoff while preserving the perceived collaborative nature of the interaction.

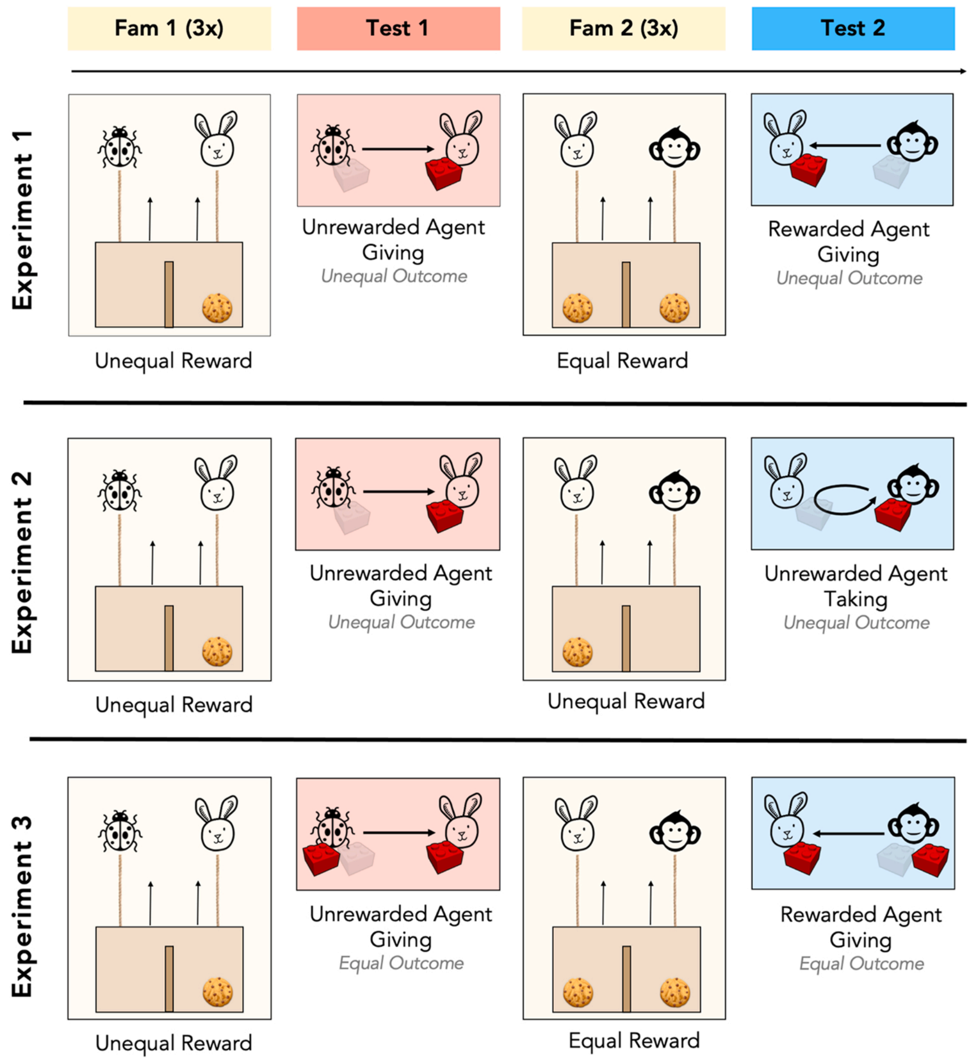

Fig. 1. Schematic visualization of the design structure and stimuli used in Experiments 1–3. Arrows in the familiarization mark the direction in which the tray was pulled. Arrows at test mark the direction of object transfer. Test events consistent with the main hypothesis are colored in red; test events inconsistent with the main hypothesis (predicted to elicit longer looking) are colored in blue. Sample stimuli are available in the OSF repository: https://osf.io/hce5p/?view_only=a0d3628538794eb2a8619e05831df45d. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2. Experiment 1

In Experiment 1 we investigated whether infants use information about reward distribution in a JA to infer the goals of the participating agents. Specifically, we tested whether infants ascribed to an agent a self-interested or altruistic goal depending on whether she reaped material rewards (or not) from her participation in the JA.

We presented infants with two types of familiarization events, each involving a different pair of agents (featuring a protagonist P acting jointly with a different partner; pair 1: P and A; pair 2: P and B). The agents worked together to pull a platform baited with either one or two cookies (Fig. 1). For one pair, the platform contained one cookie, accessible only to P, thus yielding an unequal payoff (P: 1; A: 0; "Unequal Familiarization"). For the other pair, the platform contained two cookies, one for each agent, thus yielding an equal payoff (P: 1, B: 1; "Equal Familiarization"). After each familiarization phase, infants saw a test event featuring agent A or B giving a novel object to P.

We hypothesized that infants should selectively interpret an agent's participation in a JA as altruistically motivated when her involvement cannot be justified by the obtainment of material rewards. On this basis, infants should interpret the behavior of the unrewarded agent (A), but not of the rewarded one (B), as directed to an altruistic goal. If so, we predicted infants to look longer to a novel altruistic action at test (giving) when performed by the rewarded agent (as it is inconsistent with the previously ascribed selfish goal) compared to the unrewarded one (as it is consistent with the previously ascribed prosocial goal).

2.1. Methods

2.1.1. Participants

Twenty-four infants were included in the analyses (12 females; mean age = 364 days; range = 353–378 days). Ten additional infants were excluded due to technical failure (n = 1), experimenter error (n = 2), fussiness (n = 3), crying (n = 2), and maximal looking times in both test trials (n = 2). All the participants recruited in the studies presented here were full-term infants with normal visual acuity and no declared clinical conditions. They all resided in Budapest (Hungary) and were recruited through the Hungarian birth database. The infants were brought to the lab by their caregivers, who were briefed about the nature of the study and signed an informed written consent prior to testing. All the experiments were approved by the United Ethical Review Committee for Research in Psychology (EPKEB) and conformed to the ethical rules and standards regarding psychological experimentation in Hungary.

The present sample size was determined a priori based on the fact that previous looking-time studies using similar designs and targeting infants' understanding of social interactions tend to yield large effect sizes (*Cohen's d* >*0.80,* e.g., Bas et al., 2023; Mascaro & Csibra, 2022; Rhodes et al., 2015; Wang & Henderson, 2018a;). A power analysis performed with G*Power 3.1 (Faul et al., 2007) indicated that testing 24 participants per condition would be sufficient to provide 95 % statistical power to detect a large effect size (d = 0.80, alpha =0.05) using a two-tailed paired-samples t test.

2.1.2. Stimuli

2.1.2.1. Familiarization. The familiarization events featured the following elements: three puppets (a rabbit – the protagonist, henceforth P; a monkey, A; and a ladybug, B), a cardboard platform with a divider in between and two ropes attached on each side, and a variable number of cookies (one or two) placed on the platform. Infants were shown two types of familiarization: Equal Familiarization and Unequal Familiarization, which differed with respect to the number of cookies present on the tray and the identity of the agent who acted jointly with P (i.e., A or B). In the Unequal Familiarization, there was only one cookie on the side of the platform operated by P, while the side operated by A remained empty; in the Equal Familiarization, there were two cookies, one on each side of the platform.

Both types of familiarization started with the two puppets facing the infant, on the two opposite sides of the stage, and the platform in the middle, with the two ropes laid in front. The puppets began to act after the infant was given a short exposure to the scene (1 s). Namely, the puppets turned towards each other and swayed a little (2.5 s). Afterwards, they synchronously performed three jumps towards the center of the stage, and each stopped in front of a rope (4 s). Each puppet grabbed the rope in front of them and started pulling. The pulling was performed synchronously four times and resulted in bringing the platform in the proximity of the puppets (17 s). At the end of the last pull, the two familiarization types diverged: in the *Unequal Familiarization* infants were shown P grabbing the single cookie present on her side of the tray and standing back up, while A stood still; in the *Equal Familiarization*, each puppet grabbed a cookie and stood up at the same time. The last frame was kept on screen for 8 s, or until the infants looked away for more than 2 consecutive seconds. The two familiarization types were equated for timing and overall duration (24 s of action and 8 s of still frame).

The event featured the following sound cues: a swooshing sound when the puppets turned towards each other; a bouncing sound for each of the jumps; a high-pitched "heave-ho" vocalization each time the puppets pulled the rope; and a bright "ta-da" sound when the puppet(s) grabbed the cookie.

2.1.2.2. Test. There were two test events featuring puppet A or B respectively giving a star-shaped block to P. The puppet that obtained a cookie alongside P in the Equal familiarization was labeled "Rewarded Agent", whereas the puppet who remained empty-handed at the end of the Unequal familiarization was labeled "Unrewarded Agent". The puppet acting as the Giver (A or B) was always the same who pulled the tray together with P in the familiarization preceding each test.

Both tests had the same structure (for 18 s total running time). They started with two puppets (P and the Giver, A or B) standing at the opposite sides of the stage, facing the infant, with a plastic object next to the Giver. At this point, the following sequence unfolded: the puppets turned towards each other (1.5 s); the Giver turned back, grabbed the object, lifted it, and turned again towards P (3.6 s); the Giver hopped six times towards P, and after the last jump, they finally passed the object to P (7.5 s); the two puppets remained facing each other for 1 s, after which the Giver turned away and hopped back to her original location in six jumps. As soon as the Giver reached her spot, both turned towards the infant in synchrony. The last frame of the test video was kept until infants looked away for 2 s or 60 s elapsed.

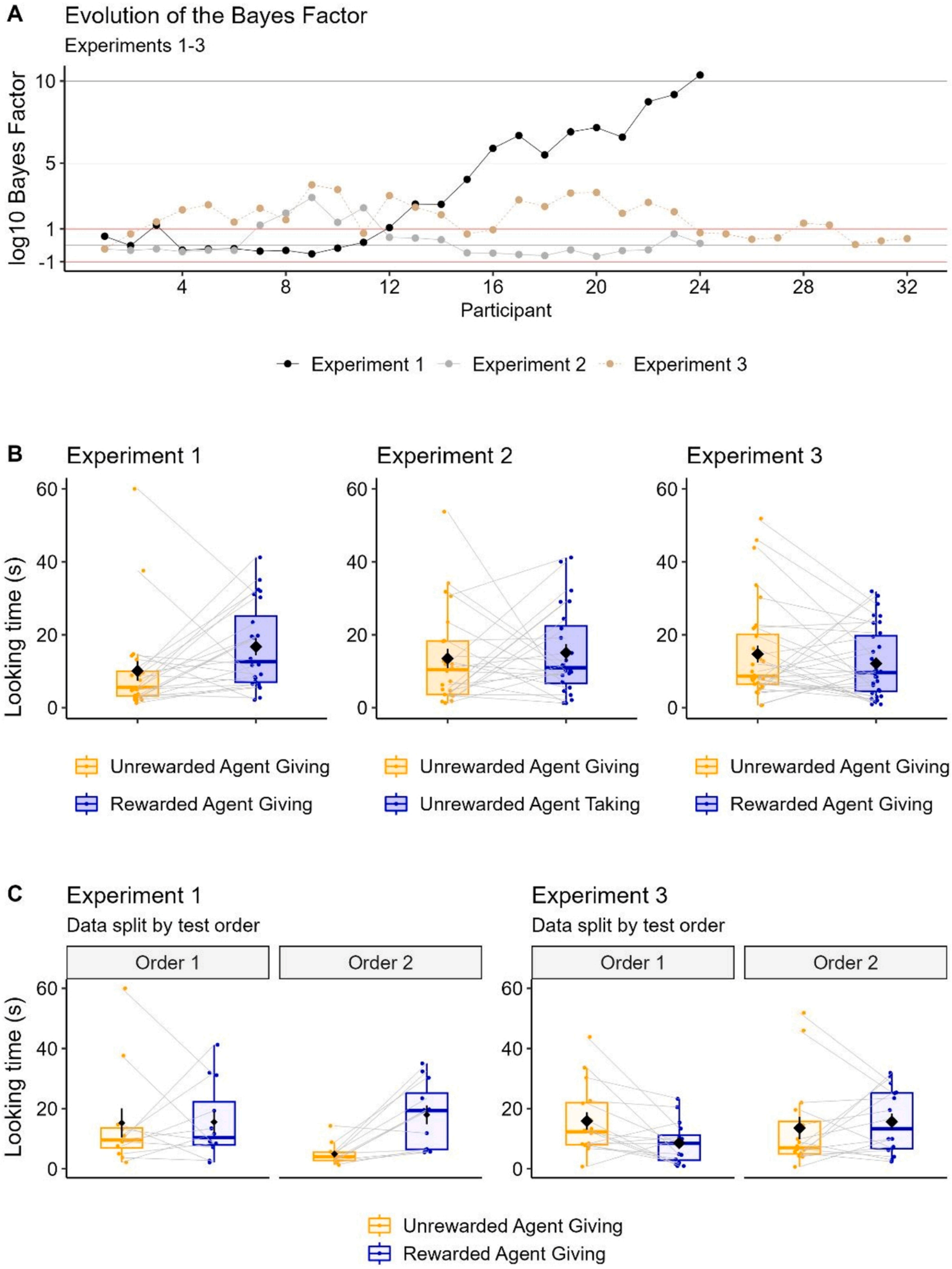

Fig. 2. Results of Experiments 1–3. (A) Evolution of log10-Bayes Factor over the course of data collection in Exps. 1–3. Values larger than + 1 indicate a strong effect in the predicted direction. (B) Raw looking times in Exps. 1–3 split by Test Type. (C) Raw looking times in Exps. 1 and 3 split by Test Type and Order (Order 1: Unrewarded Agent Giving First; Order 2: Rewarded Agent Giving First). (B, C) Black diamonds represent means and error bars represent + /- 1 standard error of the mean. The bottom and the top of the boxes represent the first and the third quartiles. Whiskers extend from the middle quartiles to the smallest and largest values within 1.5 times the interquartile range. Points connected across boxes represent individual data points.

Apart from the identity of the Giver, the overall kinematics, timing, and duration of the actions in the two test events were identical. The location of the puppets at test corresponded with the one they respectively occupied during the preceding familiarization (e.g., if P pulled the platform from the left during familiarization, it occupied the left side of the stage at test).

The event featured the following sound cues: a swooshing sound when the puppets turned towards each other; a bright sound when the Giver lifted her object; a bouncing sound when the Giver moved towards and away from P; and a different bright sound when the object was transferred to P.

2.1.3. Design and procedure

Infants were presented with 2 experimental blocks, for a total of 8 trials. Each block consisted of 3 familiarization trials (either Equal or Unequal) and 1 test trial (Rewarded Agent Giving following the Equal Familiarization and Unrewarded Agent Giving following the Unequal Familiarization), and featured one pair of puppets (P and A; P and B). Puppet P was located on one side of the stage in block 1, and on the opposite side in block 2. All trials were preceded by a short attention-getter: a green, red, or yellow star looming on a white background accompanied by a spring sound (2 s). Infants' looking behavior was coded on-line by a trained experimenter who indicated whether the infant looked at the screen.

The following factors were fully counterbalanced across infants: block order (Equal Familiarization first vs. Equal Familiarization second); identity of P's partner (puppet A as Rewarded Agent vs. puppet B as Rewarded Agent); side of P in the first familiarization block (left vs. right).

The infants were tested in a dimly lit soundproof room. They sat on the parent's lap, 100 cm away from the presentation screen (a 102 cm wide-screen LCD monitor). A hidden camera mounted above the screen recorded infants' looking behavior at 29 frames per second. The parent was asked to close their eyes, hold the infant by the waist, and avoid interacting with them during testing.

The same design and procedure were used across all experiments.

2.1.4. Data analysis

We performed an off-line frame-by-frame coding of infants' looking behavior. To be included in the final data analysis, infants had to: (1) attend for at least 50 % of each familiarization trial, from the beginning of the event to its completion (26 s); and (2) attend for at least 50 % of each test event from the beginning of the event to its completion (18 s). Looking times during familiarization were measured from the beginning of the event to when the infant looked away for more than 2 s (consecutively) or 8 s elapsed after the event's completion. Looking times at test were measured from the end of the event to when the infant looked away for more than 2 s (consecutively) or 60 s elapsed. Because infants' looking times were not normally distributed, we log-transformed the raw data before analysis to approximate a normal distribution, following the recommendation by Csibra and colleagues (2016). For ease of reading, we reported untransformed looking times in Fig. 2.

We calculated a $\log_{10}$-Bayes Factor ($\log_{10}$-BF) assuming variable effect size (after Csibra et al., 2016). We compared a null model to an alternative model that assumed change in looking times between conditions. In addition, we conducted frequentist statistical analyses.

For each experiment, fifty percent of the sample was randomly selected and re-coded by a second coder blind to the experimental hypothesis. The inter-coder agreement was excellent, as indicated by the high intra-class correlation coefficient for both consistency (Experiment 1: $r = 0.994$, Experiment 2: $r = 0.996$, Experiment 3: $r = 0.991$) and absolute agreement (Experiment 1: $r = 0.995$, Experiment 2: $r = 0.995$, Experiment 3: $r = 0.991$).

In all experiments, infants attended to the two types of familiarization events at comparably high rates: Experiment 1 (Equal: $M = 30.86$, $SD = 2.76$; Unequal: $M = 32.46$, $SD = 2.19$); Experiment 2 (Unequal featuring Agent A: $M = 32.53$, $SD = 1.04$; Unequal featuring Agent B: $M = 31.42$, $SD = 1.92$); Experiment 3 (Equal: $M = 31.66$, $SD = 2.08$; Unequal: $M = 33.04$, $SD = 1.17$).

2.2. Results & discussion

Infants looked significantly longer to the test involving the Rewarded Agent Giving ($M = 16.74$, $SD = 11.60$) than to the test involving the Unrewarded Agent Giving ($M = 10.09$, $SD = 12.98$), providing very strong evidence that they discriminated between the two events, $\log_{10}$-BF $= 10.367$; $t(23) = 3.003$, $p = .006$, $d = 0.542$ (Fig. 2). Seventeen of 24 infants showed this looking pattern ($p = .030$, by a Wilcoxon signed-ranks test). To explore whether infants' looking was affected by the order of experimental conditions, we performed an ANOVA with Test Type (Rewarded Agent Giving vs. Unrewarded Agent Giving) as a within-subject factor and Test Order (Rewarded Agent Giving first vs. Rewarded Agent Giving second) as a between-subject factor. This is analysis produced two significant effects: a main effect of Test Type, $F(1, 22) = 12.392$, $p = .002$, $\eta_p^2 = .36$, and an interaction between Test Type and Test Order, $F(1, 22) = 9.594$, $p = .005$, $\eta_p^2 = .30$. This interaction was due to the fact that infants who received the Rewarded Agent Giving test first looked significantly longer to this test event, $t(11) = 6.111$, $p < .001$, $d = 1.611$, while infants who received it second looked similarly to the two test events, $t(11) = 0.251$, $p = .806$, $d = 0.19$. This kind of interaction effect is common in looking-time studies and has been interpreted as an additive effect of the test condition and the tendency to look longer in earlier trials (e.g., Baillargeon, 1987; Csibra et al., 1999; Mascaro & Csibra, 2012).

These results are consistent with our hypothesis that an agent's participation in a JA in the absence of material rewards led infants to interpret their behavior as directed to the goal of helping the partner (in obtaining her reward). Since the two familiarization types (Equal vs. Unequal) were equated for coordination cues (e.g., the agents greeting each other and operating the tray simultaneously), these factors alone cannot account for the adoption of a prosocial interpretation of the agent's behavior in the Unequal Familiarization event. This evidence suggests that infants exploited payoff information to disambiguate the goal of the participating agents: depending on whether the agents obtained a material reward or not, infants inferred that their participation in the JA was motivated by self-interested or altruistic goals, respectively, as evinced by the different expectations that infants held about these agents performing a novel altruistic action (giving) at test.

Nonetheless, alternative explanations for this looking-time pattern are possible. For instance, infants may have looked less at the Unrewarded Agent Giving test simply because the overall number or distribution of objects (0−1) was similar across familiarization and test events. By this account, infants may not have found the Unrewarded Agent Giving test consistent with the preceding familiarization because the two events were similarly indicative of a prosocial disposition, but rather because they both culminated in asymmetric reward distributions.

3. Experiment 2

In Experiment 2 we attempted to tease apart two alternative explanations that may account for the observed pattern of results in Experiment 1: namely, whether infants similarly interpreted the agent's behavior in the unequal-familiarization JA and at test as prosocially motivated, or whether they merely reacted to the fact that both events featured a similar number (or distribution) of resources. To this end, we presented a new infant sample with two Unequal Familiarizations involving two pairs of agents (pair 1: P obtained one cookie, A none; pair 2: P obtained one cookie, B none). At test, infants were shown agent A giving an object to P, and agent B taking an identical object from P. If infants interpreted the participation of the unrewarded agents (A or B) in the JA as directed to helping P, they should look longer upon seeing agent B taking, insofar as this cost-imposing action would contradict the altruistic construal of the agent's behavior primed during familiarization. Conversely, if infants only reacted to changes in the overall resource number (or distribution) from familiarization to test, they should look equally to the two test events, despite the different valence and causal structure of the test actions, insofar as both preserved the same asymmetric payoff of the Unequal Familiarization.

3.1. Methods

3.1.1. Participants

As in Experiment 1, the final sample size consisted of 24 infants (12 females; mean age = 369 days; range = 357–377 days). Three additional infants were tested but excluded from the analyses due to experimenter error (n = 1) and fussiness (n = 2).

3.1.2. Stimuli and design

As in Experiment 1, the task consisted of two blocks, each containing 3 familiarization trials immediately followed by a test trial. Unlike in Experiment 1, both blocks featured Unequal Familiarization events: i.e., the agent pulling the tray with P always remained empty-handed. At test, infants were shown an Unrewarded Agent *Giving* test, used in Experiment 1, and a novel Unrewarded Agent *Taking* test. The Unrewarded Agent *Taking* test unfolded as follows: the two puppets turned towards each other; P turned back, grabbed the object, lifted it, and turned again towards the Taker; at this point, the Taker hopped six times towards P, and took the object from her after the last hop; the two puppets briefly faced each other, after which the Taker turned away and hopped back to her original location; as soon as the Taker reached her destination, she turned towards the infant together with P. The overall kinematics, timing, and duration of the actions across Unrewarded Puppet Taking and Giving test events were identical (18 s total running time).

3.2. Results & discussion

Contrary to our original hypothesis, infants did not look significantly longer to the Unrewarded Agent Taking test ($M = 15.08$, $SD = 11.90$) compared to the Unrewarded Agent Giving test ($M = 13.50$, $SD = 13.04$), $t(23) = 0.610$, $p = .548$, $d = 0.13$. We obtained a $\log_{10}$-BF of 0.119, providing only anecdotal evidence in favor of our hypothesis, and considered inconclusive (Biel & Friedrich, 2018). Only half of the infants looked longer at the Unrewarded Agent Giving events ($p = .511$ by a Wilcoxon-signed rank test). An ANOVA including Order (Unrewarded Agent Giving first v. Unrewarded Agent Giving second) as a between-subject factor and Test (Unrewarded Agent Giving v. Unrewarded Agent Taking) as a within-subject factor yielded no significant effects, $ps > 0.550$.

The results of Experiment 2 failed to support the hypothesis that the absence of material incentives induced the interpretation of an agent's participation in a JA as prosocially motivated. Were this the case, infants should have looked longer upon seeing the agent taking away her partner's possession, but they did not. This interpretation rests however on the assumption that infants should have interpreted taking as a cost-imposing action, and thus inconsistent with an altruistic disposition. It is possible that our taking test event failed to induce such an interpretation. Indeed, in order to ensure that the action sequence of the giving and taking test were closely matched for timing and kinematics, we introduced a number of cues that may have primed an interpretation of taking as an affiliative form of transfer: P and the Taker looked at each other; P held up the object as if offering it to the Taker and patiently waited for the Taker to seize it; and the object transfer occurred without resistance on the part of the object's original possessor. Unlike the taking actions used in other studies (cf. Gazes et al., 2017; Margoni et al., 2018), whose cost-imposing nature was made manifest through forced expropriation, the action used in Experiment 2 resembled the behavioral choreography of "facilitated taking", which one-year-olds have been shown to interpret as indicative of communal-sharing relations (Tatone & Csibra, 2020). Under this reading, the infants' comparable looking at the two test events may have then reflected the fact that both test actions were consistent with the valence of the previously ascribed prosocial goal in familiarization: where giving closely reproduced the cost-benefit relation underlying the JA event at the individual level, with A incurring costs to benefit P, taking reversed it, while being nevertheless compatible with either a principle of reciprocation (i.e., P returned the favor by letting B acquire her resource) or with a communal-sharing logic.

Caveats notwithstanding, these results lend support to the possibility that infants in Experiment 1 merely responded to the asymmetric reward distribution that the Unequal Familiarization and the Giving test event culminated in, irrespective of the action observed. Nonetheless, the lack of looking-time difference across conditions suggests that infants did not respond to changes in the agents' *individual* payoffs, otherwise they would have reacted to the taking test (as it featured a previously unrewarded agent acquiring a resource). Instead, the present findings support the possibility that infants may have reacted to changes in the overall number or distribution of resources between the two experimental phases.

4. Experiment 3

The aim of Experiment 3 was to directly test the hypothesis that infants only reacted to changes in reward distribution from familiarization to test. To this end, we presented a new infant sample with the same familiarization as in Experiment 1 and slightly modified our test stimuli: the Giver now possessed two objects instead of one. Due to this change, both test events culminated in an equal outcome (with each agent having one object). If infants in Experiment 1 responded to changes in reward distribution, then Experiment 3 should produce the opposite looking-time pattern: infants should look longer at the Unrewarded Agent Giving test insofar as the payoff that this event resulted in (1−1) is different from the one in the Unequal Familiarization (0−1). Conversely, if infants responded to the conceptual consistency of the giving action with the previously ascribed prosocial action, the same looking behavior should obtain in Experiment 3: i.e., infants should look longer to the test event featuring the rewarded agent giving compared to the unrewarded one.

4.1. Methods

4.1.1. Participants

Thirty-two infants participated in the experiment (16 females; mean age = 364 days; range = 352–375 days). Twelve additional infants were excluded from the analyses due to fussiness (n = 6), crying (n = 2), inattentiveness during test (n = 2), and experiment error (n = 2).

In light of the inconclusive results of Experiments 1–2, we decided to increase the sample size to 32 and use a pre-registered stopping rule stating that the data collection will stop either after reaching the predetermined sample size or when the sequential $\log_{10}$-Bayes Factor ($\log_{10}$-BF) in the looking-time analysis assuming variable effect size (Csibra et al., 2016) becomes larger than + 1 or smaller than − 1. The $\log_{10}$-BF calculation was fixed to be performed first after collecting 16 valid samples, and at every 8th sample thereafter (because 3 two-level factors were counterbalanced in the current design). The $\log_{10}$-BF value larger than + 1 would indicate a strong effect in the predicted or to the opposite direction, and the value smaller than − 1 would indicate an absence of looking-time difference. The minimum and maximum sample sizes, of 16 and 32 respectively, were selected based on a meta-analysis by Csibra and colleagues (2016), providing that testing 32 participants in a within-subject design has been typically sufficient to demonstrate a looking-time difference between conditions with 0.95 probability (alpha =0.05), while testing 16 participants has been sufficient to demonstrate an effect with 0.75 probability (alpha =0.05).

Determination of sample size, predictions, and statistical analyses were pre-registered at: https://osf.io/9mdwc?view_only= 2ce0d9e8a1bb4a8c819976b9e5cda674.

4.1.2. Stimuli

Infants were shown the same familiarization movies used in Experiment 1. At test, they were presented with modified test movies featuring two identical objects in front of the Giver, one which was never acted on, while the other was transferred to P. Overall action kinematics, timing, and duration were otherwise identical to the test events used in Experiment 1.

4.2. Results & discussion

Infants' looking behavior was similar across both test events (Rewarded Agent Giving test: $M = 12.12$, $SD = 9.49$; Unrewarded Agent Giving test: $M = 14.72$, $SD = 13.24$), $t(31) = 0.815$, $p = .421$, $d = 0.225$, with the BF analysis providing only anecdotal, thus inconclusive, evidence in favor of our hypothesis, $\log_{10}$-BF = 0.410. At the individual level, 19 out of 32 infants looked longer to the Unrewarded Puppet Giving test ($p = .246$, by Wilcoxon signed-ranks test). Note, however, that an ANOVA with Order (Unrewarded Agent Giving first v. Rewarded Agent Giving first) as a between-subject factor and Test Type (Rewarded Agent Giving vs. Unrewarded Agent Giving) produced a significant interaction between Test Order and Test Type, $F(1, 30) = 5.698$, $p = .023$, $\eta_p^2 = .16$, other $ps > 0.389$. Upon exploring this interaction, we found that infants who saw the Unrewarded Agent Giving first tended to look longer to this event, $t(15) = 2.110$, $p = .052$, $d = 0.77$, while those who saw the Rewarded Agent Giving first showed no differential response to the two tests, $t(15) = 1.191$, $p = .252$, $d = 0.16$.

These results show that infants tended to look longer to the test event that involved a change in the number or distribution of resources from familiarization (1−0−1−1), albeit this effect was mediated by an interaction with order. Just like in Experiment 1, we documented two additive tendencies governing infants' looking behavior: a tendency to look longer at surprising events, and a tendency to look longer at the event that was presented first (as in: Baillargeon, 1987; Csibra et al., 1999; Mascaro & Csibra, 2012). Overall, these results suggest that the findings of Experiment 1 may have been due to the infants reacting to changes in the number or distribution of rewards across the two experimental phases. This account adequately explains why in Experiment 1, where the two types of changes (goal and reward number/distribution) covaried within the same event (Rewarded Agent Giving), infants looked longer at this event, whereas in Experiment 3, where the two types of changes occurred in different test events, infants looked longer to the test featuring a change of reward number/distribution (Unrewarded Agent Giving). Corroborating this interpretation, a comparison of the infants' looking behavior between Experiment 1 and 3 by means of a repeated-measures ANOVA with test trial as within-subjects factor (Rewarded Agent Giving vs. Unrewarded Agent Giving) and Experiment (1 vs. 3) as between-subjects factor revealed a significant interaction ($F(1, 54) = 5.590$, $p = .022$, $\eta^2_p = .094$), but no main effects, other $ps > 0.305$.

5. General discussion

The present study investigated whether infants use payoff information to disambiguate the goals of agents involved in a joint action (JA). We hypothesized that when an agent reaps material rewards from a JA, her participation can be readily interpreted as directed to the obtainment of these rewards, thereby inducing the ascription of a self-interested goal. Conversely, when an agent does not obtain such rewards, further goal hypotheses should be entertained to make sense of her behavior. One such hypothesis, made salient by the collaborative nature of the event, is that the agent participated in the JA to assist the partner, thus leading to the ascription of a prosocial goal.

We examined this hypothesis across three looking-time experiments. We familiarized infants to JA events involving two participants (a common protagonist P working together with a different partner: agent A or agent B), which resulted in either the acquisition of a resource for each participant (1:1) or only for P (1:0), and tested whether the involvement of an agent in the JA despite the lack of material rewards prompted infants to interpret her behavior as prosocially motivated. Supporting our hypothesis, in Experiment 1 infants looked less to an altruistic test action (giving) directed to the former JA partner when this action was performed by a previously unrewarded agent. This looking-time pattern was consistent with the claim that infants interpreted the behavior of the unrewarded agent as altruistically motivated and thus compatible with the occurrence of a prosocial behavior at test. However, follow-up experiments failed to corroborate this interpretation. In Experiment 2, infants did not react differentially across test events involving an unrewarded agent performing either a prosocial action (giving) or an antisocial one (taking), which should be inconsistent with an altruistic interpretation of the agent's action in familiarization. Furthermore, in a close replication of the original study (Experiment 3), where each test event featured a different type of change from familiarization (i.e., number/distribution of resources or goal), infants looked longer to the test featuring the first type of change (Unrewarded Agent giving), albeit the action performed by this agent should have matched her previously ascribed goal.

Together, these findings suggest that infants' looking behavior may have been primarily driven by changes in payoff structure, such as the number and/or distribution of resources, from familiarization to test. This evidence indicates that infants encoded the payoff information of the observed JA events but failed to use this information to draw inferences about the goals of the agents involved.

It is unclear why and under which description infants represented the payoff information of the JA events. We discuss three alternatives, arranged in ascending representational complexity. First, infants may have encoded the number of objects presented at familiarization (e.g., one vs. two), irrespective of the social context these were embedded in, due to the objects being attentionally highlighted via their displacement. If this account is correct, then the findings of Experiment 1 should replicate even if the objects are made salient via a non-social manipulation (e.g., self displacemet). The rich experimental evidence showing that infants attend to object numerosity, if under subitizing range (e.g., Van Loosbroek & Smitsman, 1990; Xu, 2003; among others), may be taken as supporting this interpretation.

A second proposal is that the objects were encoded because they represented the agents' possessions. Under this reading, infants in familiarization registered whether one or both agents acquired a resource and reacted to test outcomes featuring a change in the number of possession relations ascribed (e.g., from two to one). This suggestion rests on the evidence that one-year-olds encode possession relations on the basis of proximity cues or the active manipulation of objects (e.g., giving or taking: Tatone et al., 2015; Tatone, Hernik, & Csibra, 2021). The findings of Experiment 2 however suggest that, while infants may have monitored the overall number of possession relations, they did not encode the specific identity of the agents standing in these relations. Were this the case, infants should have reacted to the taking test, because it featured reversal of the induced possession relations: the agent who acquired an object in the familiarization event ended up empty-handed at test, and vice versa for her partner. Unlike the first account, this proposal entails that infants required evidence of goal-directed acquisitive behaviors to encode the objects (necessary for representing them as the agents' possessions), but regardless of whether these were obtained through a JA. From this follows that Experiment 1 should replicate even if the objects are obtained via non-collaborative means (e.g., by having each agent pulling a separate tray).

A final possibility is that infants may have interpreted the JA in familiarization as producing equal gains for the two agents irrespective of the observed reward distribution (equal vs. unequal) but entailing different costs when the number of action steps performed by the two agents did not match (as in the unequal familiarization case). If infants interpreted this cost-reward discrepancy using a retributive logic, they may have then generated expectations about appropriate effort compensation at test. This account rests on several assumptions, made plausible by previous studies, which we unpack below.

First, contra our original hypothesis, it suggests that infants may default on assuming that agents who engage in a collaborative activity always stand to gain from its outcome even when the interaction culminates with only one of them gaining access to a reward (as in the unequal familiarization event). This idea is supported by the finding that, after being familiarized with an agent who retrieves a resource with the help of a partner, infants expect them to share it at test (Wang & Henderson, 2018a; Vorobyova, 2021). Had infants expected P to share the cookie with the unrewarded agent at the end of the tray-pulling interaction (regardless of its resolution), they would not have interpreted the two familiarization types (equal vs. unequal) as differing in the rewards the agents eventually derived. Instead, they may have perceived these events as differing in the costs that the two agents incurred to bring about the JA outcome, since in the unequal familiarization event P performed the extra action step of picking the object from the tray (while her partner stood still). If infants detected an asymmetry between (equal) benefits and (unequal) costs in familiarization, they may evaluated the resulting distribution at test following a retributive logic. Supporting this possibility is a fast-growing literature suggesting that infants' expectations of reward allotment are influenced by differences in efforts invested in a collaborative interaction (e.g., slacking off vs. contributing equally to a shared goal: Sloane et al., 2012; Wang & Henderson, 2018b). Under this account, the results obtained can be then explained as follows: infants in Experiment 1 were surprised by seeing agent P being the only agent rewarded at test, if P paid the same costs as her partner in the JA (because both agents deserved equal compensation); conversely, infants in Experiment 3 were surprised by seeing P and her partner obtaining the same reward amount at test, if P had previously worked more in the JA (because she was entitled to a higher reward).

Noteworthily, this interpretation supposes that the giving action featured at test, despite being identical across experiments (i.e., A or B transfers one resource to P), was imbued with a different meaning depending on the donor's pre-existing endowment. In Experiment 1, where the donor gave her only possession to P, the action was interpreted as benefitting the partner; conversely, in Experiment 3, where the donor gave one of her two possessions to P, the action was interpreted as equalizing rewards. Originally, we assumed that infants would interpret giving as an altruistic action irrespective of this contextual information, but in doing so we overlooked the possibility that actively splitting endowments may not be taken as a unilateral prosocial action, but as an instance of egalitarian sharing (cf. Wang & Henderson, 2018a). Since distributive expectations at test are expected to be calibrated on the agents' prior investment in the JA, this account (unlike the previous two) posits a further requirement for the encoding of reward distribution in familiarization: namely, that the activity bringing it about is perceived as collaborative. Further, the account also assumes these expectations to be agent-specific; replacing one of the agents at test with a novel character should then suffice to abolish them.

To sum up, above we offered three theoretically motivated explanations of why infants encoded what we referred to as "payoff information", and what this information may amount to in each case (e.g., number of objects, possession relations). Future studies should be able to arbitrate among these accounts, since each generates falsifiable predictions, which we spelled out above. Absent further evidence, however, parsimony considerations prompt us to consider the first account as the most plausible, as it posits minimal background assumptions (namely, that the highlighting of an object through its displacement would promote its encoding). Before taking a broader perspective on some additional research avenues inspired by our research question, we discuss below two points of possible concern in our design.

We tested 12-month-olds because, by the end of the first year of life, infants appear to have a robust understanding of the actions used in the study: giving and taking (Tatone et al., 2015), reaching an object by pulling the surface it sits on (Sommerville & Woodward, 2005), and bringing about a shared outcome by means of synchronous and similar actions (Begus et al., 2020). Critically, however, infants at this age may struggle to set up embedded goal representations (cf. Woo & Spelke, 2020), required to interpret the behavior of the unrewarded agent as altruistic (i.e., agent A's goal is to reduce the costs of P's goal completion). It is therefore conceivable that, even if the lack of material rewards may have encouraged infants to search for alternative goal hypotheses, representational constraints prevented them from appropriately setting up the second-order goal of helping.

Further, our hypothesis presupposed infants to be able to detect the similarity of (prosocial) goals across radically different interaction types: a joint action with asymmetric payoffs in familiarization and a unilateral social action at test (giving). This capacity would require a concept of prosociality sufficiently abstract to be applied across disparate instantiations. The only evidence to date directly pertaining to this issue suggests that infants draw links between helping and distributive fairness by 14 months of age (Surian et al., 2018) and generalize instrumental helping, if provided with multiple behavioral instances, at 17 months of age (Duh et al., 2022). Considering that these results come from older infants, it is plausible that the ability of relating different behavioral exemplars based on their common social orientation may not yet be developed in 12-month-olds.

A different way to approach our research question while avoiding these design pitfalls would be to test whether differences in reward distribution modulate the encoding of the participants' identity in a JA pair. Prior research suggest infants appeal to a principle of "explanatory parsimony" when interpreting behaviors occurring in facultatively interactive settings: e.g., they omit patients from the representation of taking, but not giving events, presumably because in the former case the action can be interpreted as directed to the goal of resource acquisition without having to consider further effects on the original resource possessor (Tatone et al., 2015; for evidence in adults: Yin et al., 2020; Yin et al., 2022). Transposing this line of reasoning in our study, we could reformulate our original hypothesis as follows: while collaboration leading to symmetric gains (mutualism) allows infants to explain away the two agents acting together as an ad-hoc arrangement aimed to facilitate the acquisition of individual rewards, collaboration leading to asymmetric gains (altruism) requires postulating prospective rewards (for instance, future reciprocation by the rewarded JA partner) that allow the unrewarded agent to recoup her participation costs. If so, infants should then form a stronger representation of a JA pair when exposed to a collaboration leading to unequal gains than to equal ones (since the former benefit distribution primes the representation of a long-term relation), and this should be reflected in a higher sensitivity to identity changes in pair composition. We are presently designing studies to test this very hypothesis.

In conclusion, the present work offered a novel explanation of why children are prone to attend to the payoff of collaborative interactions. We hypothesized that they may use this information not only to form expectations about remuneration based on the agents' investment in a JA ("should A be rewarded for her efforts?": Sloane et al., 2012), but also to disambiguate the types of individual goals supporting an agent's participation in the JA in the first place ("did A join forces for his own gain or to help B?"). The current experiments did not yield support for our original hypothesis in 12-month-olds, possibly because of the design shortcomings discussed above, or because the age group selected was not appropriate for the task at hand. Nevertheless, from a developmental viewpoint it is worth highlighting that the issue of how to arbitrate among different interpretations of an agent's behavior in interactive settings shares important similarities with later-developing attributional processes such as the so-called "overjustification effect". This phenomenon describes the tendency of external rewards to diminish the intrinsic motivation to engage in costly actions (for a review: Tang & Hall, 1995; for evidence in children: Warneken & Tomasello, 2014; 2008). A leading explanation for this effect is that material rewards undermine the reputational benefits that derive from engaging in socially virtuous conducts because these induce audiences to explain an agent's prosocial behavior as incentivized by personal benefits (cf. Ariely et al., 2009). Broadly considered, the overjustification model and our account both posit that observers leverage payoff information to reason about agents' goals in interactive settings. For this reason, notwithstanding their critical difference (one is a model of motive attribution, in which material rewards call into question the diagnostic value of prosocial actions; the other is a model of goal attribution, where material rewards exhaust goal search by providing an immediate rationale for an agent's behavior), we see promise in considering the possibility that mechanisms of teleological reasoning envisaged by our account may bootstrap the development of attributional processes, which require an understanding of the reputational consequences of social behavior (e.g., Heyman et al., 2007; Silver & Shaw, 2018).

Funding

This work was supported by an Advanced Investigator Grant of the European Research Council (#742231, PARTNERS).

CRediT authorship contribution statement

Denis Tatone: Conceptualization, Methodology, Data analysis, Investigation, Writing. **Laura Schlingloff:** Conceptualization, Methodology, Editing. **Barbara Pomiechowska:** Conceptualization, Methodology, Data analysis, Editing.

Data availability

Data and materials are available in the following OSF repository: https://osf.io/hce5p/?view_only=a0d3628538794eb2a8619e05831df45d

Acknowledgements

We thank Dorottya Mészégető, Petra Karman, and Iuli Savos for helping with stimuli preparation and infant testing, as well as Maria Toth and Agnes Volein for overseeing infant recruitment. Further, we thank Gergely Csibra for the insightful discussions that eventually inspired this study.

References

Ariely, D., Bracha, A., & Meier, S. (2009). Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. *American Economic Review, 99*(1), 544–555.
Baillargeon, R. (1987). Young infants' reasoning about the physical and spatial properties of a hidden object. *Cognitive Development, 2*(3), 179–200.
Bas, J., Sebastian-Galles, N., Csibra, G., & Mascaro, O. (2023). Infants' representation of asymmetric social influence. *Journal of Experimental Child Psychology, 226*, Article 105564.
Begus, K., Curioni, A., Knoblich, G., & Gergely, G. (2020). Infants understand collaboration: Neural evidence for 9-month-olds' attribution of shared goals to coordinated joint actions. *Social Neuroscience, 15*(6), 655–667.
Csibra, G., & Gergely, G. (2007). 'Obsessed with goals': Functions and mechanisms of teleological interpretation of actions in humans. *Acta psychologica, 124*(1), 60–78.
Csibra, G., Gergely, G., Bíró, S., Koos, O., & Brockbank, M. (1999). Goal attribution without agency cues: The perception of 'pure reason' in infancy. *Cognition, 72*(3), 237–267.
Csibra, G., Hernik, M., Mascaro, O., Tatone, D., & Lengyel, M. (2016). Statistical treatment of looking-time data. *Developmental Psychology, 52*(4), 521.
Duh, S., Goldman, E. J., & Wang, S. H. (2022). The role of intentionality in infants' prediction of helping and hindering. *Journal of Cognition and Development*, 1–24.
Eason, A. E., Doctor, D., Chang, E., Kushnir, T., & Sommerville, J. A. (2018). The choice is yours: Infants' expectations about an agent's future behavior based on taking and receiving actions. *Developmental Psychology, 54*(5), 829.
Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191.
Fawcett, C., & Gredebäck, G. (2013). Infants use social context to bind actions into a collaborative sequence. *Developmental Science, 16*(6), 841–849.
Fawcett, C., & Gredebäck, G. (2015). Eighteen-month-olds, but not 14–month-olds, use social context to bind action sequences. *Infancy, 20*(1), 115–125.
Gazes, R. P., Hampton, R. R., & Lourenco, S. F. (2017). Transitive inference of social dominance by human infants. *Developmental Science, 20*(2), Article e12367.
Hamlin, J. K., & Wynn, K. (2011). Young infants prefer prosocial to antisocial others. *Cognitive Development, 26*(1), 30–39.
Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature, 450*(7169), 557–559.
Henderson, A. M., & Woodward, A. L. (2011). "Let's work together": What do infants understand about collaborative goals? *Cognition, 121*(1), 12–21.
Heyman, G. D., Fu, G., & Lee, K. (2007). Evaluating claims people make about themselves: The development of skepticism. *Child Development, 78*(2), 367–375.
Margoni, F., Baillargeon, R., & Surian, L. (2018). Infants distinguish between leaders and bullies. *Proceedings of the National Academy of Sciences, 115*(38), E8835–E8843.
Mascaro, O., & Csibra, G. (2022). Infants expect agents to minimize the collective cost of collaborative actions. *Scientific Reports, 12*(1), 1–8.
Mascaro, O., & Csibra, G. (2012). Representation of stable social dominance relations by human infants. *Proceedings of the National Academy of Sciences USA, 109*(18), 6862–6867.
Rhodes, M., Hetherington, C., Brink, K., & Wellman, H. M. (2015). Infants' use of social partnerships to predict behavior. *Developmental Science, 18*(6), 909–916.
Silver, I. M., & Shaw, A. (2018). Pint-sized public relations: The development of reputation management. *Trends in Cognitive Sciences, 22*(4), 277–279.
Sloane, S., Baillargeon, R., & Premack, D. (2012). Do infants have a sense of fairness? *Psychological Science, 23*(2), 196–204.
Sommerville, J. A., & Woodward, A. L. (2005). Pulling out the intentional structure of action: the relation between action processing and action production in infancy. *Cognition, 95*(1), 1–30.
Surian, L., Ueno, M., Itakura, S., & Meristo, M. (2018). Do infants attribute moral traits? Fourteen-month-olds' expectations of fairness are affected by agents' antisocial actions. *Frontiers in Psychology, 9*, 1649.
Tang, S. H., & Hall, V. C. (1995). The overjustification effect: A meta-analysis. *Applied Cognitive Psychology, 9*(5), 365–404.

Tatone, D. & Csibra, G. Infants infer different types of social relations from giving and taking actions (2020). In S., Denison, M. L., Mack, M. L., Y. Xu, & B. Armstrong (Eds.) *Proceedings of the forty second annual conference of the cognitive science society* (pp. 2981–2987).
Tatone, D., Geraci, A., & Csibra, G. (2015). Giving and taking: Representational building blocks of active resource-transfer events in human infants. *Cognition, 137*, 47–62.
Van Loosbroek, E., & Smitsman, A. W. (1990). Visual perception of numerosity in infancy. *Developmental Psychology, 26*(6), 916.
Vorobyova, E. (2021). *Infants' understanding of cooperation* (PhD dissertation). Retrieved from ⟨https://www.etd.ceu.edu/2021/vorobyova_elizaveta.pdf⟩.
Xu, F. (2003). Numerosity discrimination in infants: Evidence for two systems of representations. *Cognition, 89*(1), B15–B25.
Yin, J., Csibra, G., & Tatone, D. (2022). Structural asymmetries in the representation of giving and taking events. *Cognition, 229*, Article 105248.
Yin, J., Tatone, D., & Csibra, G. (2020). Giving, but not taking, actions are spontaneously represented as social interactions: Evidence from modulation of lower alpha oscillations. *Neuropsychologia, 139*, Article 107363.
Wang, Y., & Henderson, A. M. (2018a). We cooperated so… now what? Infants expect cooperative partners to share resources. *Infant Behavior and Development, 52*, 9–13.
Wang, Y., & Henderson, A. M. (2018b). Just rewards: 17-month-old infants expect agents to take resources according to the principles of distributive justice. *Journal of Experimental Child Psychology, 172*, 25–40.
Warneken, F., & Tomasello, M. (2008). Extrinsic rewards undermine altruistic tendencies in 20-month-olds. *Developmental Psychology, 44*(6), 1785.
Warneken, F., & Tomasello, M. (2014). Extrinsic rewards undermine altruistic tendencies in 20-month-olds. *Motivation Science, 1*(S), 43–48.
Woo, B., & Spelke, E.S. (2020). How to help best: Infants' changing understanding of multistep actions informs their evaluations of helping. In S., Denison, M. L., Mack, M. L., Y. Xu, & B. Armstrong (Eds.) *Proceedings of the forty second annual conference of the cognitive science society* (pp. 384–390).