COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 50 (2026) e70167

ISSN: 1551-6709 online
DOI: 10.1111/cogs.70167

No Evidence for Agent–Patient Role Attribution in Human Infants, Human Adults, and Guinea Baboons (*Papio papio*)

Floor Meewis,[a,b] Iris Barezzi,[a] Marielle Hababou-Bernson,[a] Joël Fagot,[a,b] Nicolas Claidière,[a,b] Isabelle Dautriche[a]

[a] *Centre de Recherche en Psychologie et Neurosciences, UMR7077, Aix-Marseille University, CNRS*
[b] *Station de Primatologie-Celphedia, CNRS UAR846*

Received 16 September 2024; received in revised form 26 October 2025; accepted 30 November 2025

Abstract

Languages describe “who is doing what to whom” by distinguishing the event roles of agent (doer) and patient (undergoer), but it is debated whether they result from nonlinguistic representations that may already exist in preverbal infants and nonhuman animals. The phenomenon of causal perception, where the subsequent movements of two objects A and B evoke the impression of A launching B, is a simple depiction of an agent–patient relation. The seminal study by Leslie and Keeble from 1987 proposed that infants of 6 months old may be able to attribute agent and patient roles to such causal displays, after they demonstrated the infants’ dishabituation upon seeing a launching event that was reversed. They introduced the idea that a role reversal had taken place upon reversing the direction of the launching event (launcher becoming launchee), but not in a noncausal temporal gap event where the agent and patient roles were not present. The present study tested this hypothesis in three different populations: 6-month-old human infants, human adults, and Guinea baboons (*Papio papio*). For the human infants, we applied a habituation-dishabituation design, and for the human adults and baboons, a conditional match-to-sample task. We were unable to replicate the findings of Leslie and Keeble in human infants. Similarly, we did not find evidence for an effect specific to reversing launching events in human adults and baboons. Inconsistent results across different

---

Correspondence should be sent to Floor Meewis, Centre de Recherche en Psychologie et Neurosciences, Faculté des Sciences, Aix-Marseille Université, 3, place Victor Hugo, 13003 Marseille, France. E-mail: floormeewis@gmail.com

studies call into question the role reversal paradigm for Michottean launches to study event role attribution.

Keywords: Event role; Causal perception; Michotte; Replication; Role reversal; Event cognition

---

1. Introduction

When communicating about the world, people tend to describe *who is doing what to whom*. Most languages achieve this by distinguishing between the doer (agent) and the undergoer (patient) of the action, typically as the subject and object of a sentence. This distinction is often highlighted through word order or case marking (Bickel, Witzlack-Makarevich, Choudhary, Schlesewsky, & Bornkessel-Schlesewsky, 2015; Dryer, 2013; Napoli & Sutton-Spence, 2014). The origin of these thematic roles—whether they arise from abstract cognitive representations that precede language—remains a frequently debated topic (Dowty, 1991; Ferretti, McRae, & Hatherell, 2001; Fillmore, 1968; Rissman & Majid, 2019; Strickland, 2017; Wilson, Zuberbühler, & Bickel, 2022; Zuberbühler, 2019, 2020, 2022; Zuberbühler & Bickel, 2022).

Insights into whether the mind represents thematic roles in the absence of language come from studies with human adults on nonlinguistic tasks. There is now rich evidence that adults are experts in extracting roles from perceptual input alone: they can grasp agent and patient roles from very brief visual scenes (Dobel, Gumnior, Bölte, & Zwitserlood, 2007; Hafri, Papafragou, & Trueswell, 2013) and do so unconsciously (Hafri, Trueswell, & Strickland, 2018). Additionally, it has become clear that adults discriminate agent from patient by displaying differential attentional behavior for these two roles. Across a wide range of modalities, whether gestures, events, or visual narratives, agents attract more visual attention than patients (Cohn & Paczynski, 2013; Isasi-Isasmendi et al., 2023; Webb, Knott, & MacAskill, 2010). Agents are also easier to identify than patients (Dobel et al., 2007; Isasi-Isasmendi et al., 2023; Wilson, Papafragou, Bunger, & Trueswell, 2011; but see: Ünal, Wilson, Trueswell, & Papafragou, 2024) and are placed first when described with nonlinguistic devices, such as gestures (Goldin-Meadow, So, Özyürek, & Mylander, 2008). While these studies make it clear that the adult mind has nonlinguistic representations of agent and patient, they cannot tell if language knowledge is necessary for these representations to exist.

Studies with human infants and nonhuman animals can reveal whether conceptual knowledge about event roles is in place before full command of natural language or even without any command of language at all. In these populations, investigations into role assignment have often been intertwined with research on causal perception. In these studies, participants are presented with 2D visual displays of geometric shapes in which an object A moves toward a stationary object B, and B starts moving as soon as A stops and touches B (see Fig. 1a), which human adults perceive as A *causing* B to move (Michotte, 1963; for review: Scholl & Tremoulet, 2000). This impression of causality is thought to be perceptually processed (Kominsky & Scholl, 2020; Moors, Wagemans, & de-Wit, 2017; Rolfs, Dambacher, & Cavanagh, 2013). When a temporal gap and/or a spatial gap is inserted in between the

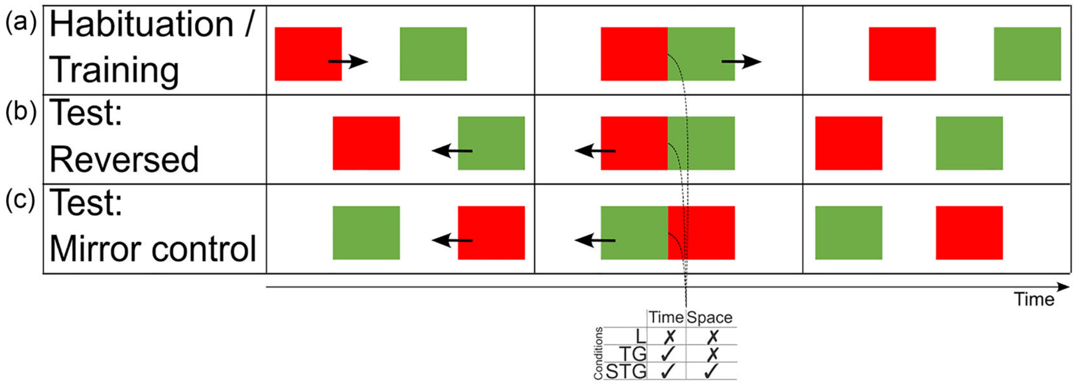

Fig. 1. Schematical examples of the Michottean events used as stimuli. During a launching event (L), there is no temporal gap (Time = ×) and no spatial gap (Space = ×) added between the movements of the objects, resulting in an event that is perceived as causal by human adults. During a temporal gap event (TG), there is a temporal gap (Time = √), but no spatial gap (Space = ×) added between the movements of the objects, resulting in a display perceived as noncausal. During a spatial and temporal gap event (STG), there is a temporal gap (Time = √) and a spatial gap (Space = √) added between the movements of the two objects, which is perceived as a noncausal relation. (a) Events used in the habituation phase for the human infants and for training for the baboons and human adults. (b) Reversed events used in the test phase. (c) Mirror control events used in the test phase.

movements of the two objects (see Fig. 1a), such a causal impression is not evoked. From birth, infants look more toward causal launches than toward noncausal events (Mascalzoni, Regolin, Vallortigara, & Simion, 2013), showing that they differentiate launching from other, similar events. From 6 months old onward, infants view direct launching events as categorically distinct from events with a temporal gap and with a spatial gap in between the objects' movements (Cohen & Amsel, 1998; Oakes, 1994). However, even though these studies show a sensitivity to the spatiotemporal properties related to causality and simultaneously suggest sensitivity to causality at this early age, they do not provide evidence that infants can represent the agent and patient roles in such displays. Similarly, nonhuman animals can discriminate between different Michottean events (newly hatched chicks [*Gallus gallus*]: Mascalzoni, Regolin, & Vallortigara, 2010; Guinea baboons [*Papio papio*]: Meewis, Barezzi, Fagot, Claidière, & Dautriche, 2024; pet dogs [*Canis familiaris*]: Völter & Huber, 2021; pigeons [*Columba livia*]: Young, Beckmann, & Wasserman, 2006) and some may even perceive similar types of events as causal (chimpanzees [*Pan troglodytes*]: Matsuno & Tomonaga, 2017). However, whether nonhumans also attribute agent and patient roles to such events is unknown. A recent study that tested whether different ape species have a bias toward agents in videos depicting real-life events showed variable results that were largely dependent on animacy cues (Brocard, Wilson, Berton, Zuberbühler, & Bickel, 2024). In sum, while the perception of causality has been examined through Michottean launches in human adults, infants, and nonhuman animals, merely being sensitive to causality or its spatiotemporal properties does not indicate whether agent and patient roles are recognized.

In infancy, the capacity to represent agent and patient roles has been argued to be present at 6 months of age (Leslie & Keeble, 1987). In this study, Leslie and Keeble (1987) aimed to test whether Michottean launching events are perceived holistically or as having an internal structure. The authors habituated one group of infants to a launching event (A launching B) and another group to a similar event with a temporal gap (A moves, stays still, then B moves), with both events involving movement from left to right. In the test phase, infants were shown the same events but in reverse temporal order: B started to move first, followed by A, with motion from the right to the left side of the screen (see Fig. 1b). Both groups of infants were thus habituated to a left to right movement and then tested on right to left movement. The authors predicted that if the infants perceive the events as causal, they may find the reversal of the launching event most surprising, as it involved a change in spatiotemporal properties as well as a change in causal direction. Indeed, the reversal of the launching event resulted in more surprise for 6-month-old infants than the reversal of the temporal gap event, albeit with substantial interindividual variability. While confirming that the infants treat the launching event as different from the temporal gap event at this age, the authors additionally interpreted the stronger dishabituation for the launching event as being caused by a role reversal. By reversing the launch, the launcher (agent) had become the launchee (patient) and the other way around, whereas such a role reversal does not occur when reversing the temporal gap event, since this noncausal event does not involve agent and patient roles. Subsequently, this study has been widely cited as evidence of sensitivity to agent−patient roles in 6-month-old infants (Kominsky & Carey, 2024; Mascalzoni et al., 2013; Muentener & Carey, 2010; Ünal, Ji, & Papafragou, 2021). However, it is important to note that this conclusion was developed to explain the observed pattern of results, rather than being directly validated by the findings. Critically, a direct replication study conducted by Bélanger and Desrochers (2001) showed looking recovery not only for the reversal of the launching event but also of the temporal gap event. They consequently failed to demonstrate significantly longer looking for the reversal of the launching event compared to the temporal gap event. However, they did show that the surprisal for the reversed launching event was significantly greater than an additional condition in which they reversed a spatial gap event (another type of noncausal event). This is why this study is often cited as a successful replication of Leslie and Keeble's effect (e.g., Kominsky & Carey, 2024; Muentener & Bonawitz, 2017) even though the original comparison (reversal of launching vs. reversal of temporal gap event) did not generate the same results. Another recent study applied the role reversal method to asymmetrical Michottean events in 8-month-old infants (Kominsky & Carey, 2024). Such an asymmetrical launching event, in which the speed of A is three times larger than the speed of B, is not perceived as different from a symmetrical launching event (Kominsky et al., 2017). During the test condition, the infants who were habituated to asymmetrical Michottean events did not look longer at the reversal of the launching compared to the temporal gap event (Kominsky & Carey, 2024), indicating that reversing these highly similar events again does not lead to the same findings as in Leslie and Keeble's study. In sum, while it is widely accepted that 6-month-old infants represent agent and patient roles while viewing Michottean events, the evidence supporting this conclusion is limited.

Beyond the role-reversal method, research on whether infants represent agent−patient roles either lack appropriate controls (such as a noncausal condition: Golinkoff, 1975; Rochat, Striano, & Morgan, 2004) or show results inconsistent with agent−patient role attribution based on their role in a launching event (but see Luo et al., 2009, for evidence that it could be the case at 5 months, in the special case where the agent is depicted by a human hand—an agent that infants have considerable experience with). For example, Cohen and Oakes (1993) investigated 10-to-12-month-old infants' reaction upon replacing the toys portraying object A and B in a launching and in a temporal gap event. The infants showed dishabituation for a change in object A and not B, regardless of whether it was a launching or a temporal gap event, thus not supporting the idea that infants assign agent and patient roles to a launching event. Similarly, in a different study, infants of 10 months old who were habituated to a launching event between two boxes covered with fur did not look longer at the agent compared to the patient when they were revealed to be hollow (Kominsky, Li, & Carey, 2022). These findings indicate that the infants did not have differential expectations about the animacy status (cued by the insides) of the two objects in the launching event and did not attribute the movement of the patient as being caused by the agent. Another study suggests that this ability may perhaps appear later in life. After habituating infants to a 2D animated launching event, 14-month-olds (but not 10-month-olds) were surprised to see a test event in which the patient engaged in a self-propelled motion but not the agent, which was not the case for the objects for the temporal gap event (Cicchino, Aslin, & Rakison, 2011). This study shows that infants made inferences about the agentive abilities of the objects based on the role they fulfilled in the events, but only at 14 months. Thus, beyond the role-reversal design, evidence suggests that, at least at 10 months, infants do not attribute agency to an object solely based on its role in a launching event. This suggests that, since the seminal study of Leslie and Keeble (1987), there has been limited empirical support for the widely held view that 6-month-old infants can represent agent and patient roles based purely on causal structure.

The present study thus aimed to replicate and extend the seminal study by Leslie and Keeble (1987). This is important for several reasons:

1. Leslie and Keeble (1987) is widely cited as evidence for early representations of the thematic roles of agent and patient.
2. The theory posited by the original study was developed post-hoc as an interpretation of the observed results, and, therefore, requires confirmatory hypothesis testing.
3. Subsequent studies by Bélanger and Desrochers (2001) and Kominsky and Carey (2024) resulted in inconsistent findings using the same or highly similar events.
4. These failed replications raise questions about the robustness of the role reversal method in detecting agent and patient representations.

In this study, we tested agent and patient role attribution to Michottean launching events using the role reversal method. We extended the original study into two key aspects.

First, we added a control condition. Previous studies have shown that even newborn infants treat launching events differently from temporal gap events (Mascalzoni et al., 2013). Therefore, it is possible that the longer looking times observed by Leslie and Keeble might be due to differences between the two events rather than specifically the role reversal. To address this, we implemented a "mirror control condition" where a reversal of direction occurs while the order of movements (roles in the launch) remains the same (see Fig. 1c). This aims to determine whether the role reversal is indeed responsible for the previously reported longer looking times. Second, we tested not only 6-month-old infants, as in the original study, but also human adults and Guinea baboons (*P. papio*). Studying human adults for whom it is well-established that they view Michottean events in terms of causality and quickly attribute roles to actors in interactions will further validate the role reversal method for studying agent and patient role attribution. Additionally, findings in nonhumans could provide insights into whether event role attribution is shared with nonhuman species, as has been recently proposed (Wilson et al., 2022).

For the infants, we implemented a habituation study replicating Leslie and Keeble's original study. Six-month-old infants were shown a launching event or a temporal gap event during habituation and were subsequently tested on the reversed and mirrored control versions during the test phase. We measured the looking time at the test events. For the human adults and the baboons, we used a conditional match-to-sample (MTS) classification task, which was designed to test whether the participants would perceive the Michottean launches in terms of causality (Meewis et al., 2024). During this task, the participants were presented with a video of one of the Michottean events, which they had to classify into one of two categories: launching event or spatial and temporal gap event. The two categories were represented by two distinct response buttons, and during training, participants had to learn the appropriate response through trial and error. After learning the categories, the participants went into the test phase, in which they were presented with the reversed and mirrored control versions of the training events. We measured the accuracy of sorting the test events.

Since the assignment of agent−patient roles appears to be a cognitive mechanism that operates rapidly and automatically (e.g., Hafri et al., 2013, 2018) with its origin in core cognition (Strickland, 2017) and/or beyond the human lineage (Wilson et al., 2022), we preregistered the expectation that the reversal of the causal launch would be perceived as different from the reversal of the noncausal event and different from the mirrored versions of the causal and noncausal events as follows from Leslie and Keeble's proposed role reversal theory (Leslie & Keeble, 1987). We expected that this distinctiveness of the role reversal should lead to a longer looking time for the infants and more classification mistakes for the human adults and baboons (cf. Davidoff & Fagot, 2010; Vauclair, Fagot, & Hopkins, 1993; where accuracy depends on how close the match is to the sample).

2. Methods

The methods presented in this article have been preregistered. The preregistration for the infant study can be found at the following link: https://osf.io/rga9f/files/5rtgb; for the human adults at the following link: https://osf.io/vt5g8/files/gcewk (see step 3); and for the baboons at the following link: https://osf.io/vt5g8/files/grcxb (see step 3). All materials have been grouped together in the repository of the infant preregistration and can thus be found at the following link: https://osf.io/rga9f.

2.1. *Ethics statement*

2.1.1. *Human infants*

The study on infants has received approval from the Comité De Protection Des Personnes Sud Meditérranee III (#2021.05.11quatre_21.05081.210511). Parents signed an informed consent for their child's participation. They received compensation for travel expenses and were gifted an "infant researcher diploma" and a baby body suit.

2.1.2. *Baboons*

The study on baboons was carried out in accordance with French and EU standards and received approval from the French Ministère de l'Education Nationale et de la Recherche (#APAFIS-2717-2015111708173794-V3). Procedures used in the present study were also consistent with the guidelines of the Association for the Study of Animal Behaviour.

2.1.3. *Human adults*

The study on human adults has received approval from the CEEI of INSERM (#20-733). We provided information about the task to the human participants, and they consented to participate voluntarily in the study after having read the instructions.

2.2. *Participants*

2.2.1. *Human infants*

Data of 53 healthy full-term infants (31 female; mean age: 28.0 ± 0.3 (SEM) weeks old, age range: 24–34 weeks) were used for the analyses; 27 were habituated to the launching event (of which 13 were tested with mirrored first and 14 with reversed first) and 26 to the temporal gap event (of which 13 were tested with mirrored first and 13 with reversed first). A total of 68 infants were tested, but 15 infants were excluded because they were preterm (4), did not habituate in under 18 trials (7), or were fussy during the experiment (4). The sample size was chosen based on Leslie and Keeble's original study, which tested 12 infants per condition, and the recommendations for habituation-dishabituation designs by Oakes (2017).

2.2.2. *Baboons*

We analyzed data of 11 Guinea baboons (*P. papio*) (10 female; mean age: 11.1 ± 1.3 (SEM) years old, age range: 5.8−17.4 years) of which six participated two times (see below). The task was made available to a total of 23 individuals who were housed at the Station de Primatologie in Rousset-sur-Arc (France) in two groups of 18 and five individuals. The baboons were tested with the automatic learning device for monkeys (ALDM) developed by Fagot and Paleressompoulle (2009). With this system, the baboons had *ad libitum* access to testing booths with touch screens where they participated voluntarily in computerized tasks employing an operant conditioning method. We were unable to use data from 12 baboons because they did not participate in the task and/or did not meet the learning criteria. The data we present here are part of a series of experiments with Michottean events that took place in May and July 2022, of which other parts have been described elsewhere (Meewis et al., 2024).

2.2.3. Human adults

We analyzed the data from 39 participants (gender: 21 female, 18 male; mean age = 27.5 ± 1.4 (SEM) years old, age range = 20–59 years) recruited through Prolific. We chose participants who reported to be fluent in English, such that they could understand the instructions. We excluded one additional participant because they did not complete the experiment in one go, but paused for several hours in the middle. Data collection took place at the end of May 2023, and this experiment was also part of a series with Michottean events, of which other parts have been described elsewhere (Meewis et al., 2024). Participants were paid £5 for their time.

2.3. Stimuli

2.3.1. Stimuli types

The stimuli used in this study are based on Michotte's launching paradigm (Michotte, 1963). These are simple 2D animations in which an object A moves first from one side to the middle of the screen, and an object B moves second from the middle to the other side of the screen. In the *launching event (L)*, A moves toward B, which immediately starts moving in the same direction with the same speed upon being touched by A, see Fig. 1a. Contrasting the L event, we used either a *temporal gap event (TG)* in which A moves toward B until they touch, then they both stay still, and B starts moving in the same direction with the same speed as A did, or *a spatial and temporal gap event (STG)* in which A moves toward B, stops before reaching B, they both stay still, after which B starts moving in the same direction with the same speed as A did, see Fig. 1a. Since L, but not TG and STG, is perceived as causal by human adults (Scholl & Tremoulet, 2000), only the L has an agent−patient structure, and the TG and STG have two separately moving agents (Leslie & Keeble, 1987; Mascalzoni et al., 2010).

In the test versions of the events, the movements of the object are reversed (starting from the opposite side of the screen). In the *reversed* event, besides the directional change, first object B moves and then object A, see Fig. 1b. In the *mirror control* version, the directional change takes place, but the object A is still the first to move followed by object B, see Fig. 1c. Importantly, a role reversal takes place only in the reversed version of the L event: when B launches A, the launcher has become the launchee, thus reversing the original agent−patient roles. In all the other conditions, either the agent−patient structure remains (mirror control L event), or there are no agent−patient roles to be reversed (reversed TG/STG event and mirror control TG/STG event).

2.3.2. Infant stimuli

For the infants, we recreated the stimuli used in Leslie and Keeble (1987): an L event and a TG event. The direction of movement was from left to right during habituation (and from right to left during test). Both objects were square-shaped; object A was red and object B was green. The background was black. Leslie and Keeble reported using objects of 3.8 cm in width and 3.1 in height, and a traveling distance for both of the objects of 18.1 cm on a screen of 44×33 cm. We kept the same ratio of stimuli to background and displayed them on a 27'' screen. The different durations of the two videos were balanced by adding a few extra still frames at the beginning and end of the direct launch video. The total durations and durations of separate movements remained the same as in the original study (see Table S1). All videos were made using the animation features of Microsoft PowerPoint and subsequently exported in mp4 format. We used a resolution of 2880×2160 pixels and a framerate of 60 frames/s (24 frames/s in Leslie and Keeble).

2.3.3. *Baboon and human adult stimuli*

We used the same stimuli for baboons and human adults: L and STG. In contrast to Leslie and Keeble (1987), we decided to use STG instead of TG events to maximize the difference from the L event, in order to facilitate discrimination training for the baboons. Additionally, using these two categories that were distinct in the temporal and spatial domains was necessary for other experiments in this series, as described in another publication (Meewis et al., 2024). We created two stimulus sets. In the first set, object A was an orange circle and object B was a purple circle, with a movement direction from left to right during training (and right to left during test). In the second stimulus set, object A was a blue square and object B was a yellow square as with a movement direction from right to left during training and from left to right during test. Stimuli were presented on a black background. The two sets used the same durations (see Table S2). All videos were made using the animation features of Microsoft PowerPoint and subsequently exported in mp4 format. We used videos with a resolution of 944×720 pixels and a framerate of 30 frames/s.

2.4. *Procedure and design*

2.4.1. *Human infants*

We used a habituation-dishabituation design just like Leslie and Keeble's (1987). Half of the participants were habituated to the L event, and the other half of the participants were habituated to a TG event. Following habituation, two test trials were displayed, one with the reversed and one with the mirror control version of the habituation video, in random order, see Fig. 1.

During stimulus presentation, the infant sat on their parent's lap, facing a screen. The caregiver wore opaque glasses or averted their eyes to avoid influencing the infant. During the experiment, the experimenter was situated outside the view of the infant.

The stimuli were presented with the help of Habit2, Version 2.2.1 (Oakes et al., 2019). During habituation, a loop of the event (L or TG) was played. The habituation criterion was put to a decrease of 50% of looking time based on the sliding window average of three consecutive trials compared to the average looking time of the first three trials (Oakes, 2010). The maximum number of trials for habituation was set at 18 (Leslie & Keeble, 1987). A trial was defined as looking at the screen and ended when the infant looked away for more than a second (Leslie & Keeble, 1987). Looking away for more than a second prompted the appearance of an attention getter (brightly colored bubbles with sound) on the screen. When looking was re-established, the next trial was presented.

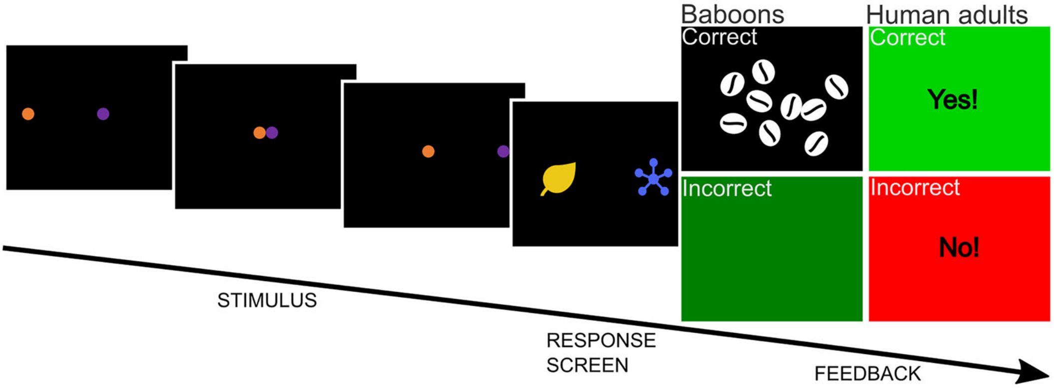

Fig. 2. Match-to-sample task for baboons and human adults. First, a video is played portraying one of the Michottean events (either a direct launch [L] or a spatial and temporal gap event [STG]). Next, a screen appears with two response buttons. For correct choices, the baboons receive a food reward and the humans see a green screen with the word “Yes!,” for incorrect choices, the baboons see a green screen for 3 s (time-out) and the humans see a red screen with the word “No!” for 3 s (time-out).

During the experiment, the infant was videotaped while looking at the videos. The experimenter used the projection of the video to determine whether the infant was looking at the screen. Both during habituation and test, looking times were recorded in real time in Habit2 by pressing a key on the keyboard of the computer. For the participants included in the analysis, the experiment took on average 2.6 min.

2.4.2. *Baboons and human adults*

The human adults and baboons participated in a conditional MTS task in which they were first trained to discriminate L from STG events and subsequently tested on the classification of the reversed and mirrored events.

Videos and response stimuli were presented to the baboons using E-prime (V2) and to the humans using Labvanced (Finger, Goeke, Diekamp, Standvoß, & König, 2017). The humans received minimal instructions before the start of the experiment. We explained that they would see videos which they had to classify by pressing one of two buttons. We also told them they would receive feedback in some of the trials (as a green or red screen), and in some other trials no feedback in the form of a purple screen. Lastly, we told them that the task’s duration would increase with wrong answers, to give them an incentive to try to answer correctly. Causality or agent−patient roles were not mentioned.

During a trial, a video displaying an L or STG event played as a sample, which was followed by two comparison stimuli of which the participant had to choose one, triggering feedback, see Fig. 2. The L event type was assigned to one comparison stimulus, and the STG event to the other, and this was counterbalanced across participants. For the first stimulus set, the comparison stimuli were a yellow leaf and a blue molecule, and for the second stimulus set, a cyan chef’s hat and a purple swan. The sample video was displayed in the center of the screen, and the comparison stimuli were displayed for the monkeys at the bottom and for the humans in the middle of the screen, with their right-left location counterbalanced across trials. For the monkeys, feedback consisted of rewarding correct responses with some wheat grains, while punishing incorrect responses with a 3 s green time-out screen. Touching the screen during the video interrupted the trial and also led to the 3 s green time-out screen. The humans received a 1 s green screen with the word “Yes!” to correct responses and were punished with a 3 s red screen with the word “No!” to incorrect responses, see Fig. 2.

The human adults and baboons underwent training to discriminate the L from STG before being subjected to the reversed and mirror control events in the test phase. In the training phase, participants first learnt to discriminate one L from one STG video. These two events were matched for their total duration (i.e., 1600 ms). Participants were presented with blocks of randomly ordered trials, such that, within a block, there were half of L trials and half of STG trials. The baboons had blocks of 60 trials, and the humans had blocks of eight trials in the training phases. To continue to the next phase of the experiment, the baboons had to reach 80% correct classification for both events within the block (60 trials), and the humans had to reach 100% correct classification for all the trials in the block (eight trials). After reaching these preset criteria, two extra events were added, one L and one STG. We intended that the participants would classify the events according to the causality of the event (although paying attention to the presence or absence of the gaps was another successful strategy). For both species, each block in the second training phase consisted of 25% of previously learnt L events, 25% of previously learnt STG events, 25% new L events, and 25% new STG events. When the baboons reached the criteria of 70% correct on all four events, they continued to the next phase. For the humans, we again placed a 100% correct criterion for all of the trials. Next, we tested whether the participants had not just memorized the four training videos and how they associated with the response stimuli, by measuring their capability to generalize to unseen L and STG events (see: Meewis et al., 2024).

Subsequently, the participants entered the test phase where we tested how baboons and humans assigned the reversed and mirror control Michottean events into these previously learnt categories. For the baboons, we used blocks of 60 trials, including eight probe trials (2 L reversed; 2 L mirror control; 2 STG reversed; 2 STG mirror control) and 52 previously learnt events (26 L and 26 STG), and repeated these blocks 16 times. The probes were randomly rewarded in 80% of cases regardless of the response given to avoid any learning for these trials. For the humans, we presented three times a block of the four probe events. We did not give the humans feedback on these probe trials, but instead showed a purple screen for 2 s after they made their choice, regardless of the response given, which was explained in the instructions as indicative of no feedback.

3. Analyses and results

3.1. Human infants

3.1.1. Analyses

All analyses were performed using RStudio (R Core Team, 2021). Since looking time data often shows a right-skewed distribution, we checked the distributions and performed logarithmic transformations where needed (Csibra, Hernik, Mascaro, Tatone, & Lengyel, 2016) following our preregistration.

First, we monitored whether there existed a preference for either of the habituation videos. To be certain that there were no differences related to habituation carrying over in the test phase, we implemented three two-way ANOVAs, one for each habituation property, checking whether there were differences regarding habituation between the habituation events (L or TG) and the test groups (Reversed first or Mirror control first) or an interaction between those. We tested as habituation properties: (1) the number of trials needed to reach the habituation criterion; (2) the mean looking time over the first three habituation trials; and (3) the mean looking over the last three habituation trials. We did not detect any differences between groups of infants habituating for the two event types (L or TG) and their following test order (order of reversed and mirror) regarding the number of trials needed for habituation (all $p > .652$), the mean duration of the first three habituation trials (all $p > .143$), and the mean duration of the last habituation trials (all $p > .134$).

Since the habituation was equal between the groups, we analyzed the test trial looking times (and not recovery time) as preregistered. For the analysis of looking times in the test trials, we used the following general linear mixed model from the lme4 package (Bates, Mächler, Bolker, & Walker, 2015): log(looking time) $\sim$ event type (L/TG) * test type (Mirror/Reversed) $+$ order $+$ (1 | participant). p-Values for main fixed effects were based on likelihood ratio tests, and simple effects were reported from the summary table of the model. Since there was no effect of order of test events on the logarithmically transformed looking times ($\chi^2 = 3.69$, $p = .055$), we removed this term from the model.

Note that nonsignificant p-values cannot differentiate between no effect (H_0) and no evidence for any conclusion. We, therefore, calculated a Bayes factor to measure the extent to which our data support a hypothesized positive interaction (H_1; see Hypothesis section below) or a negative to no interaction (H_0). Following Dienes (2014), we compute a Bayes factor modeling H_1 as a half-normal (thus testing a one-sided prediction, since we are testing a directional hypothesis) with a mean of 0 and the SD set to a rough estimate of the predicted interaction if H_1 is true. This method requires three numbers: (1) an estimated mean interaction in the data; (2) the associated standard error; and (3) an estimate of the predicted interaction under H_1. For (1) and (2), we use the estimate of the interaction and associated standard error obtained in the linear mixed model. For (3), we would base this on previous data, but there is no study that used mirror and reversed events within the same design, so we used the motivated-maximum approach (Silvey, Dienes, & Wonnacott, 2024) and set the estimated interaction as twice the value of the intercept of our model.

We added an extra, not preregistered, test to verify whether the infants dishabituated upon seeing the test events. Here, we calculated the recovery times by subtracting the looking time for the last habituation trial from the first test trials after habituation (disregarding the second test trial) and tested whether these were significantly above zero with one-sided one-sample t-tests.

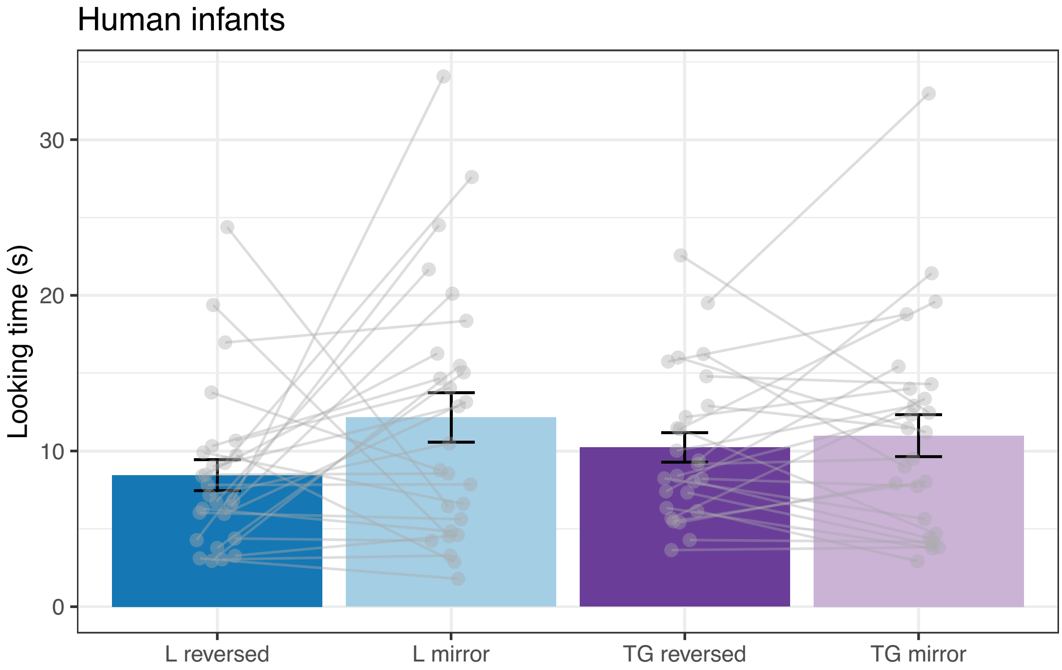

Fig. 3. The human infants' looking times during the test events L reversed, L mirror, TG reversed, and TG mirror. Error bars reflect the standard error of the mean. Gray dots and connecting lines show individual data.

3.1.2. *Hypothesis*

If infants are sensitive to agent–patient roles within the role reversal paradigm, we predict an interaction between event type (L vs. TG) and test type (Reversed vs. Mirror). Specifically, infants should exhibit greater dishabituation to the reversed launching event compared to the reversed temporal gap event. This outcome would indicate that infants detect the reversal of agent–patient roles in launching events, whereas such a role reversal is not present in temporal gap events. Conversely, no difference in dishabituation is expected between launching and temporal gap events in the mirror condition, as the agent–patient roles remain unchanged in this test type. We, therefore, expected a significant positive interaction between event type and test type.

3.1.3. *Results*

The infants looked on average 8.5 ± 1.0 (SEM) s to the reversed launching event, 12.2 ± 1.6 (SEM) s to the mirrored launching event, 10.2 ± 0.9 (SEM) s to the reversed temporal gap event, and 11.0 ± 1.4 (SEM) s to the mirrored temporal gap event, see Fig. 3. There was no interaction between event type and test type ($\chi^2 = 2.08$, $p = .149$). The Bayes factor (BF_{10}) in favor of the positive interaction was 0.22, indicating moderate evidence for H_0. There was no main effect of event type ($\chi^2 = 0.54$, $p = .461$) nor of test type ($\chi^2 = 1.79$, $p = .181$) on the infants' looking times. Thus, contrary to Leslie and Keeble's results, infants did not dishabituate more to the reversed launching event than to the reversed temporal gap event, nor to the mirrored events, suggesting no effect of the role reversal.

Using the recovery times (looking time first test trial minus looking time last habituation trial), we found that the infants dishabituated for all of the test events: the reversed launching event (2.0 ± 1.1 (SEM) s recovery; $t(13) = 1.81$, $p = .047$), the mirrored launching event (6.7 ± 2.0 (SEM) s recovery; $t(12) = 3.29$, $p = .003$), reversed temporal gap event (3.6 ± 1.4 (SEM) s recovery; $t(12) = 2.65$, $p = .011$), and the mirrored temporal gap event (6.4 ± 1.6 (SEM) s recovery; $t(12) = 4.09$, $p < .001$).

3.1.4. Analyses exact replication

From all the tested participants, we selected the participants who reflected the paradigm of Leslie and Keeble exactly. We thus picked those participants who were presented with the reversed version of the habituation trials as the first test trial. In other words, we removed all the mirrored trials and the participants who saw the mirrored version before the reversed version. This resulted in 27 infants (14 female; mean age: 28.0 ± 0.5 (SEM) weeks old, age range: 24–32 weeks) meeting these criteria, of whom 14 infants habituated to the launching event and 13 infants habituated to the temporal gap event. This sample is comparable to Leslie and Keeble's.

We calculated the recovery looking time as the looking time for the first test trial minus the looking time for the last habituation trial (cf. Leslie & Keeble, 1987). The recovery times of the reversed launching event and reversed temporal gap event were compared with a one-tailed *t*-test, as in the original study. The Bayes factor was calculated using the ttestBF function from the BayesFactor package (Morey et al., 2015).

3.1.5. Results exact replication

Using the data of the subset of infants who were presented with the reversed version of the habituation trials as the first test trial, the recovery looking time of the L group was 2.0 ± 1.1 (SEM) s and of the TG group 3.8 ± 1.4 (SEM) s, see Fig. 4. These looking times were significantly different from zero in the L group ($t(13) = 1.81$, $p = .047$) and in the TG group ($t(12) = 2.65$, $p = .011$), confirming that the infants dishabituated upon seeing the new events. There was, however, no difference in recovery time between the infants habituated to L and those to TG ($t(23.5) = -0.2$, $p = .816$). This analysis yielded a Bayes factor BF_{10} of 0.28, indicating that, given the data, there was moderate evidence for the null hypothesis (no stronger recovery for the infants who habituated to L compared to the infants who habituated to TG) to be true than not true. The results did not replicate Leslie and Keeble, see Fig. 4.

3.2. Baboons

3.2.1. Analyses

To test the response accuracy for mirrored and reversed events, we implemented a generalized linear mixed model with binomial error structure using the lme4 package (Bates

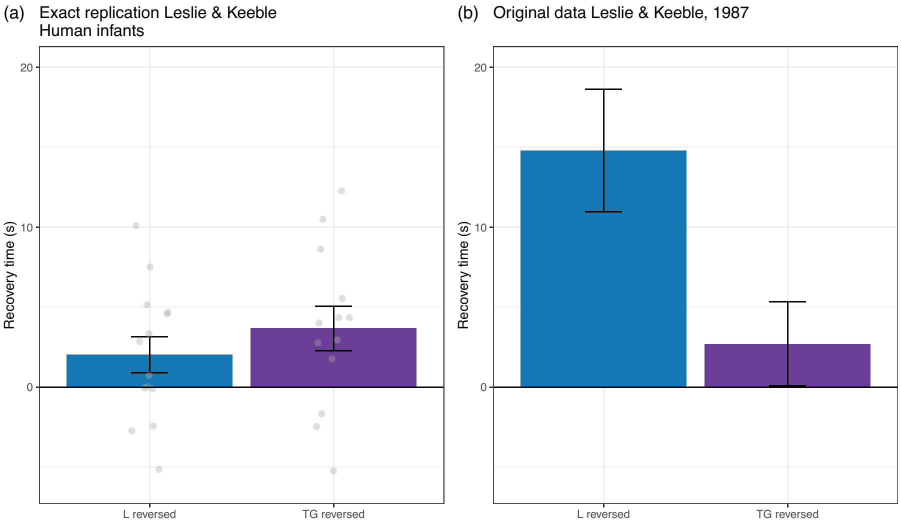

Fig. 4. The human infants' recovery looking times during the test events L reversed and TG reversed. Error bars reflect the standard error of the mean. Gray dots show individual data. (A) Data from the present study. (B) Original data reproduced from Leslie and Keeble (1987).

et al., 2015): accuracy (0/1) ~ event type (L/STG) * test type (Mirror/Reversed) + stimulus set (first/second) + (1 | participant). We kept the stimulus set in the model because it had a significant effect on the response accuracy, such that they were more accurate on the first set during the first period than on the second set during the second period (62.6 ± 3.9 (SEM) vs. 57.9 ± 3.1 (SEM)% correct; $\chi^2 = 5.41$, $p = .020$). We decided to deviate from our preregistered methods and remove the baseline (accuracy on L baseline was 73.7 ± 2.1 (SEM)% correct and accuracy on STG baseline 73.8 ± 2.3 (SEM)(% correct) from the analysis and only use the test events, since their accuracy on baseline resulted from training and thus only the accuracy on the mirrored and reversed events was of interest. See the infant analyses section for the Bayes factor calculation.

3.2.2. *Hypothesis*

Following our predictions with the infants, the role reversal should make the reversed launching event distinctively different from the launching event, while no such distinction is expected between the reversed STG event and the STG event, and between the mirror events and their nonmirrored versions.

Generally, in MTS paradigms, the more dissimilar the match is from the sample, the lower the accuracy (Vauclair et al., 1993; Davidoff & Fagot, 2010). We thus predicted that if baboons pick up event roles in Michottean launches, we should observe an interaction between test type and event type: they should be less accurate to classify the reversed launching event (as "launching") than the reversed STG event (as "STG"), while no difference in accuracy was expected between the mirror launching event and the mirror STG events.

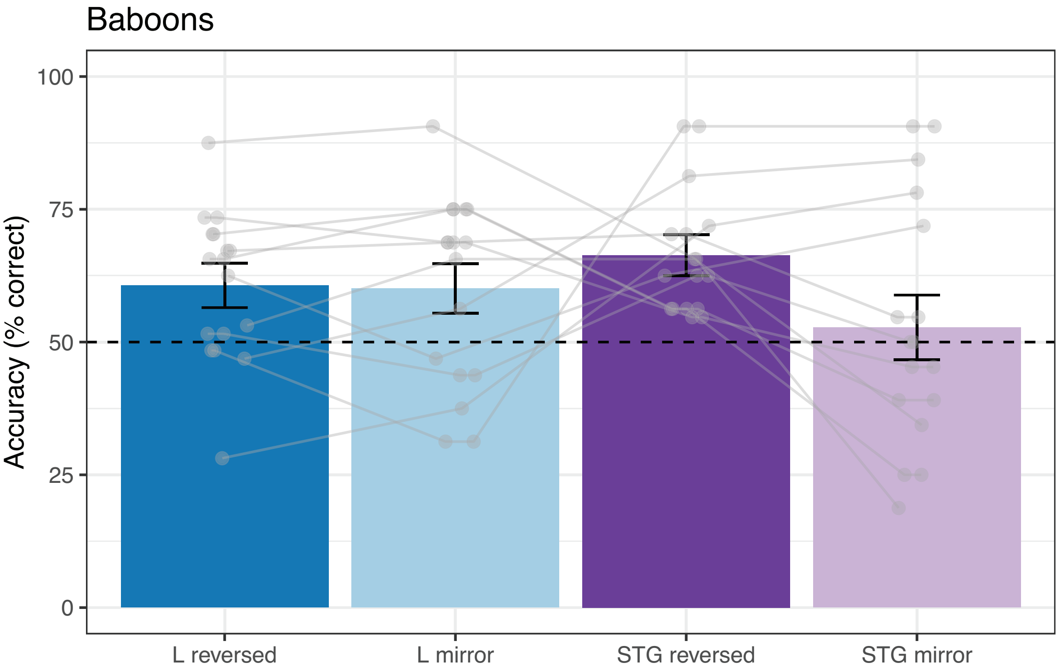

Fig. 5. The Guinea baboons' accuracy scores in percentage correct for the events L reversed, L mirror, STG reversed, and STG mirror. Error bars reflect the standard error of the mean. Gray dots and connecting lines show individual data. The black dotted line shows the chance level of 50% correct.

3.2.3. *Results*

For the baboons, we found an interaction between test type and event type ($\chi^2 = 9.72$, $p = .002$), yielding a BF_{10} of 0.038, suggesting strong evidence for H_0. There was a main effect of test type (reversed vs. mirror; $\chi^2 = 11.44$, $p < .001$) and no evidence for a main effect of event types (L vs. STG; $\chi^2 = 0.09$, $p = .76$), see Fig. 5. Yet, the interaction was not in the predicted direction: post-hoc comparisons showed that the baboons did not differ in classifying the reversed L compared to the reversed STG (60.7 ± 4.2 (SEM)% correct vs. 66.4 ± 3.9 (SEM)% correct; $z = -1.96$, $p = .050$). They were significantly better at classifying the mirrored L compared to the mirrored STG (60.1 ± 4.7 (SEM)% correct vs. 52.8 ± 6.1 (SEM)% correct; $z = 2.45$, $p = .014$). The accuracy scores for the reversed L and mirrored L did not differ from each other ($z = 0.19$, $p = .952$), whereas they were less accurate for the mirrored STG compared to the mirrored STG ($z = 4.57$, $p < .001$). In sum, the baboons did not perform worst on the reversed launching event compared to the other events, suggesting no evidence for an effect of role reversal.

3.3. *Human adults*

3.3.1. *Analyses*

To test the response accuracy for mirrored and reversed events, we implemented a generalized linear mixed model with binomial error structure using the lme4 package (Bates et al.,

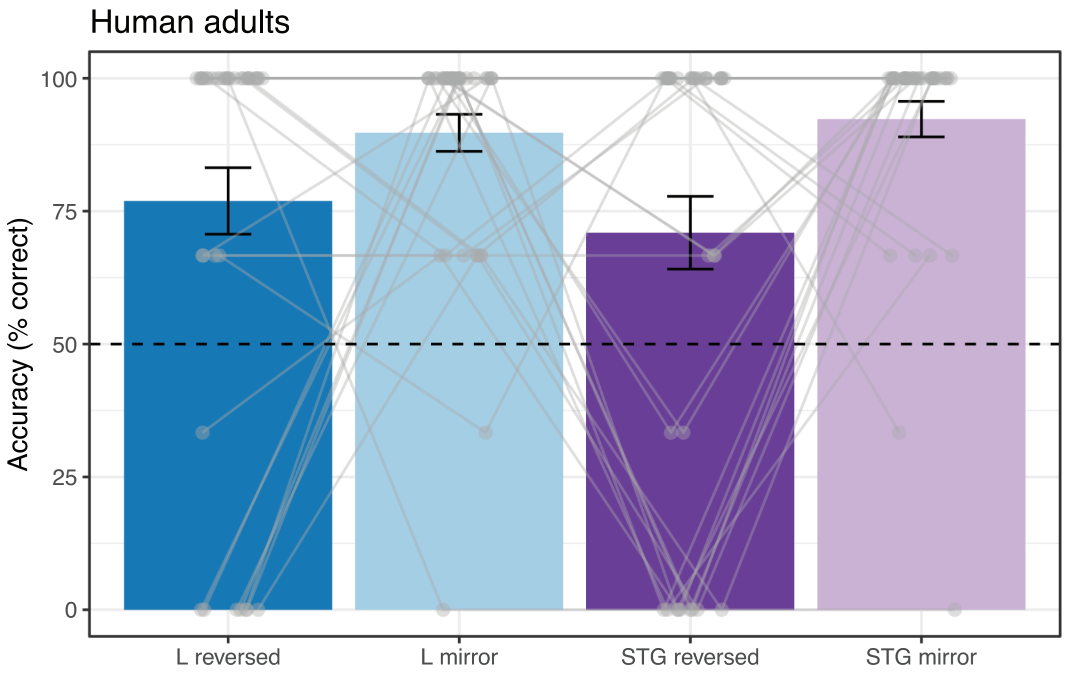

Fig. 6. The human adults' accuracy scores in percentage correct for the events L reversed, L mirror, STG reversed, and STG mirror. Error bars reflect the standard error of the mean. Gray dots and connecting lines show individual data. The black dotted line shows the chance level of 50% correct.

2015): accuracy (0/1) ~ event type (L/STG) * test type (Mirror/Reversed) + (1 | participant). As before, we computed a Bayes factor following the procedure described in the infants' analyses section.

3.3.2. *Hypothesis*

As with the baboons, if human adults are sensitive to event roles within the role reversal paradigm, we expected an interaction between test type (Mirror/Reversed) and event type (L/STG): Adults should make more mistakes for classifying the reversed L event compared to the reversed STG event, while no such performance difference is expected between the mirror L and the mirror STG events.

3.3.3. *Results*

The model on the human responses did not reveal an interaction effect between test type and event type on accuracy ($\chi^2 = 2.12$, $p = .145$). The Bayes factor (BF_{10}) was 0.034, indicating strong evidence for H_0. There was a main effect of test type, indicating that the human adults performed better at classifying the mirrored than the reversed events (91.0 ± 2.4 (SEM) vs. 73.9 ± 4.6 (SEM)% correct; $\chi^2 = 36.05$, $p < .001$), see Fig. 6. There was no effect of event type, showing that the human adults performed equally on the L and the STG events (83.3 ±

3.6 (SEM) vs. 81.6 ± 4.0 (SEM)% correct; $\chi^2 = 0.39, p = .53$). Hence, the human adults did not perform worst on the reversed launching event, showing no evidence for an effect of the role reversal.

4. Discussion

This study showed no evidence of agent and patient role attribution to Michottean launching events by means of the role reversal method in three different populations: human infants, Guinea baboons, and human adults. Importantly, we did not replicate the findings by Leslie and Keeble (1987), after which the notion that preverbal infants of 6 months represent agent and patient began gaining momentum. Our results thus cast doubt on the robustness of the role reversal paradigm and/or the assumed ability of 6-month-olds to assign agent−patient roles to causal displays.

For infants, our results are in contrast with Leslie and Keeble's study. One may argue that our study employed a slightly different procedure (inclusion of a mirror control condition leading to two test trials), which could have resulted in these null findings. However, the additional analysis focusing on the subset of infants who saw the reversed test trial first, exactly copying the method of Leslie and Keeble, shows the same pattern of results. Aside from our replication attempt, two other, more recent, studies could not replicate Leslie and Keeble's effect or the magnitude of effect (Bélanger & Desrochers, 2001; Kominsky & Carey, 2024). Altogether, three studies have resulted in inconsistent findings trying to replicate the results of Leslie and Keeble, calling into question the robustness of the original findings.

One possible explanation for this nonreplication is that 6-month-old infants may not, in fact, automatically extract agent and patient roles from Michottean launches. This would be consistent with other studies that show that infants at 10 months of age do not make attribution of agency on the sole basis of its role in a launching event (Cicchino et al., 2011; Kominsky et al., 2022). Since causality perception is assumed to be present at this age (e.g., Cohen & Amsel, 1998), this would point to a potentially problematic dichotomy between sensitivity to causality and the ability to represent agent−patient roles. One possibility is that the apparent sensitivity to causality in previous studies actually reflects a sensitivity to continuous spatiotemporal properties rather than to causality itself. Cohen and Amsel (1998) and Oakes (1994) showed that 6- and 7-month-old infants who habituated to a launching event dishabituated upon seeing a temporal gap or a spatial gap event, but infants habituated to a temporal gap event did not dishabituate to a spatial gap event (and vice versa). This result is consistent with the idea that infants group events based on perceived causality, but it is also compatible with infants grouping together events based on constant, sustained movement (both spatial and temporal events as interrupted movements). This alternative interpretation would be in line with infants not attributing roles to launching displays.

A second possibility is that 6-month-old infants are indeed sensitive to causality and do extract agent−patient roles, but the role reversal method with Michottean launches may not reliably capture this ability. This interpretation is further supported by the adult findings: human adults showed no evidence of perceiving event roles in this design, despite the well-established findings that human adults perceive Michottean events in terms of causality (Scholl & Tremoulet, 2000, see also Meewis et al., 2024, for confirmation that this is the case for this group of adults in subsequent steps of our task). Support for the idea that human infants, human adults, and even baboons are sensitive to agent—patient roles outside Michottean events comes from other implicit tasks involving chasing events (see also, for other event types Papeo et al., 2024 for infants, and Cohn & Paczynski, 2013; Dobel et al., 2007; Hafri et al., 2013, 2018; Webb et al., 2010; Wilson et al., 2011 for adults). While observing 2D animations of a triangle chasing another triangle, both human adults and 6-month-old infants prefer to look toward the agent of a chasing interaction (Galazka & Nyström, 2016; Meyerhoff, Schwan, & Huff, 2014), indicating that they discriminate between the two event roles. Similarly, in baboons, another experiment from our lab has demonstrated that baboons have an attentional preference for the agent in such chasing interaction (Meewis, Fagot, Claidière, & Dautriche, 2025), revealing that, under the right circumstances, Guinea baboons appear to discriminate agent and patient roles.

Why are infants, adults, and baboons sensitive to event roles in chasing events but not in Michottean launching events? As noted in the introduction, previous studies have failed to show that by 10 months of age, infants attribute agency based solely on an object's role in a launching event. Instead, they require additional features, such as animacy cues, to do so (Cicchino et al., 2011; Kominsky et al., 2022). Chasing events may naturally provide such cues: to depict a chase between triangles, each triangle must move in a motion pattern suggestive of animacy. Similarly, in Papeo et al. (2024), which used events involving one person acting on another, infants may readily identify the characters as potential agents. This contrasts with Michottean events, where no obvious animacy features are inferable from the objects involved.

Could the use of an explicit task with human adults (and baboons) have obscured the desired effect? Unlike the experiments on chasing events discussed above, the current set of studies with human adults and baboons employed an explicit task, in which participants first learned to categorize launching and spatiotemporal events separately. Such an explicit learning task may not be well suited to reveal the features of an event to which adults are sensitive, but rather those of which they are most consciously aware. For instance, human adults performed worse when classifying reversed events (involving two changes from the training event: direction and color) compared to mirrored events (involving only a directional change), regardless of event causality, suggesting that role reversal was less salient than the low-level perceptual changes. This suggests that low-level changes, such as direction or color change, are the most relevant features that are used by human adults in the categorization task, not the roles of the objects. To conclusively determine whether the role reversal task itself, rather than the explicit nature of the task, leads to an inability to demonstrate agent and patient role attribution, future studies should aim to replicate these findings using an implicit task in human adults. However, it is important to note that in other implicit tasks, such as visual adaptation paradigms involving launching events, human adults have been shown to be insensitive to causal roles while remaining sensitive to low-level features, such as event direction (Ohl & Rolfs, 2024), suggesting that the detection of agent–patient roles may not depend solely on whether the task is implicit or explicit.

In sum, the findings of our experiments raise doubt about the ability of 6-month-old infants to attribute agent and patient roles to Michottean launches and/or the effectiveness of the role reversal method to address this question. Our findings are consistent with previous evidence showing that infants younger than 14 months have difficulty attributing causal roles to simple shapes purely based on causal structure, unless additional cues about animacy are provided (Cicchino et al., 2011; Kominsky et al., 2022; Meewis et al., 2025). Given the lack of effect of the role reversal across all our populations, including human adults, it is also possible that there is an issue with the role reversal method itself. Consequently, the use of the role reversal method for testing the agent and patient role attribution in Michottean launches needs reconsideration. Such reconsideration may have wide-ranging consequences, as Leslie and Keeble's study is a cornerstone in the field of cognitive development and has inspired many subsequent studies over a period of more than 30 years. It thus necessitates a re-evaluation of a broad body of work applying role reversals to study agent and patient role attributions.

From this study alone, we cannot draw definitive conclusions about the existence of cognitive representations of agent and patient in the absence of language. This thus remains a crucial question with potentially far-reaching implications for the development and evolution of language. Verification of previous, influential studies is paramount to build a solid research foundation to bring us closer to understanding how causal events and their event roles are processed.

Acknowledgments

This work was supported by the Agence Nationale de la Recherche (ANR-20-CE28-0005) and the Institute for Language, Communication and the Brain (ILCB) through the Agence Nationale de la Recherche (ANR-16-CONV-0002). We thank Julie Gullstrand, Sebastien Barniaud, and the staff at the Primatology Station for their help with the experiments with the baboons.

Funding statement

This work was supported by the Agence Nationale de la Recherche (ANR-20-CE28-0005) and the Institute for Language, Communication and the Brain (ILCB) through the Agence Nationale de la Recherche (ANR-16-CONV-0002).

Conflict of interest

We have no competing interests to declare.

Data availability statement

The methods presented in this article have been preregistered. The preregistration for the infant study can be found at the following link: https://osf.io/rga9f/files/5rtgb; for the human adults at the following link: https://osf.io/vt5g8/files/gcewk (see step 3); and for the baboons at the following link: https://osf.io/vt5g8/files/grcxb (see step 3). All materials have been grouped together in the repository of the infant preregistration and can thus be found at the following link: https://osf.io/rga9f.

Ethics approval statement

The study on infants has received approval from the Comité De Protection Des Personnes Sud Meditérranee III (#2021.05.11quatre_21.05081.210511). Parents signed an informed consent for their child's participation.

The study on baboons was carried out in accordance with French and EU standards and received approval from the French Ministère de l'Education Nationale et de la Recherche (#APAFIS-2717-2015111708173794-V3). Procedures used in the present study were also consistent with the guidelines of the Association for the Study of Animal Behaviour.

The study on human adults has received approval from the CEEI of INSERM (#20-733). We provided information about the task to the human participants, and they consented to participate voluntarily in the study after having read the instructions.

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using **lme4**. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bélanger, N. D., & Desrochers, S. (2001). Can 6-month-old infants process causality in different types of causal events? *British Journal of Developmental Psychology*, *19*(1), 11–21. https://doi.org/10.1348/026151001165930

Bickel, B., Witzlack-Makarevich, A., Choudhary, K. K., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2015). The neurophysiology of language processing shapes the evolution of grammar: Evidence from case marking. *PLOS ONE*, *10*(8), e0132819. https://doi.org/10.1371/journal.pone.0132819

Brocard, S., Wilson, V. A. D., Berton, C., Zuberbühler, K., & Bickel, B. (2024). A universal preference for animate agents in hominids. *iScience*, *27*(6), 109996. https://doi.org/10.1016/j.isci.2024.109996

Cicchino, J. B., Aslin, R. N., & Rakison, D. H. (2011). Correspondences between what infants see and know about causal and self-propelled motion. *Cognition*, *118*(2), 171–192. https://doi.org/10.1016/j.cognition.2010.11.005

Cohen, L. B., & Amsel, G. (1998). Precursors to infants' perception of the causality of a simple event. *Infant Behavior and Development*, *21*(4), 713–731. https://doi.org/10.1016/s0163-6383(98)90040-6

Cohen, L. B., & Oakes, L. M. (1993). How infants perceive a simple causal event. *Developmental Psychology*, *29*(3), 421–433. https://doi.org/10.1037/0012-1649.29.3.421

Cohn, N., & Paczynski, M. (2013). Prediction, events, and the advantage of Agents: The processing of semantic roles in visual narrative. *Cognitive Psychology*, *67*(3), 73–97. https://doi.org/10.1016/j.cogpsych.2013.07.002

Csibra, G., Hernik, M., Mascaro, O., Tatone, D., & Lengyel, M. (2016). Statistical treatment of looking-time data. *Developmental Psychology*, *52*(4), 521–536. https://doi.org/10.1037/dev0000083

Davidoff, J., & Fagot, J. (2010). Cross-species assessment of the linguistic origins of color categories. *Comparative Cognition & Behavior Reviews*, *5*, 100–116. Retrieved from https://comparative-cognition-and-behavior-reviews.org/vol5_davidoff_fagot/

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in psychology*, *5*, 781.

Dobel, C., Gumnior, H., Bölte, J., & Zwitserlood, P. (2007). Describing scenes hardly seen. *Acta Psychologica*, *125*(2), 129–143. https://doi.org/10.1016/j.actpsy.2006.07.004

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, *67*(3), 547. https://doi.org/10.2307/415037

Dryer, M. S. (2013). Order of subject, object and verb. In M. S. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Zenodo. https://doi.org/10.5281/zenodo.7385533

Fagot, J., & Paleressompoulle, D. (2009). Automatic testing of cognitive performance in baboons maintained in social groups. *Behavior Research Methods*, *41*(2), 396–404. https://doi.org/10.3758/BRM.41.2.396

Ferretti, T. R., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language*, *44*(4), 516–547. https://doi.org/10.1006/jmla.2000.2728

Fillmore, C. J. (1968). Lexical entries for verbs. *Foundations of Language*, *4*(4), 373–393.

Finger, H., Goeke, C., Diekamp, D., Standvoß, K., & König, P. (2017). LabVanced: A unified JavaScript framework for online studies. In *International Conference on Computational Social Science*.

Galazka, M., & Nyström, P. (2016). Infants' preference for individual agents within chasing interactions. *Journal of Experimental Child Psychology*, *147*, 53–70. https://doi.org/10.1016/j.jecp.2016.02.010

Goldin-Meadow, S., So, W. C., Özyürek, A., & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences*, *105*(27), 9163–9168. https://doi.org/10.1073/pnas.0710060105

Golinkoff, R. M. (1975). Semantic development in infants: The concepts of agent and recipient. *Merrill-Palmer Quarterly of Behavior and Development*, *21*(3), 181–193.

Hafri, A., Papafragou, A., & Trueswell, J. C. (2013). Getting the gist of events: Recognition of two-participant actions from brief displays. *Journal of Experimental Psychology: General*, *142*(3), 880–905. https://doi.org/10.1037/a0030045.Getting

Hafri, A., Trueswell, J. C., & Strickland, B. (2018). Encoding of event roles from visual scenes is rapid, spontaneous, and interacts with higher-level visual processing. *Cognition*, *175*(February 2017), 36–52. https://doi.org/10.1016/j.cognition.2018.02.011

Isasi-Isasmendi, A., Andrews, C., Flecken, M., Laka, I., Daum, M. M., Meyer, M., Bickel, B., & Sauppe, S. (2023). The agent preference in visual event apprehension. *Open Mind*, *7*, 240–282. https://doi.org/10.1162/opmi_a_00083

Kominsky, J. F., & Carey, S. (2024). Infants' representations of Michottean triggering events. *Cognition*, *250*, 105844. https://doi.org/10.1016/j.cognition.2024.105844

Kominsky, J. F., Li, Y., & Carey, S. (2022). Infants' attributions of insides and animacy in causal interactions. *Cognitive Science*, *46*(1), e13087. https://doi.org/10.1111/cogs.13087

Kominsky, J. F., & Scholl, B. J. (2020). Retinotopic adaptation reveals distinct categories of causal perception. *Cognition*, *203*, 104339. https://doi.org/10.1016/j.cognition.2020.104339

Kominsky, J. F., Strickland, B., Wertz, A. E., Elsner, C., Wynn, K., & Keil, F. C. (2017). Categories and constraints in causal perception. *Psychological Science*, *28*(11), 1649–1662. https://doi.org/10.1177/0956797617719930

Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognitive Development*, *25*, 265–288.

Luo, Y., Kaufman, L., & Baillargeon, R. (2009). Young infants' reasoning about physical events involving inert and self-propelled objects. *Cognitive psychology*, *58*(4), 441–486.

Mascalzoni, E., Regolin, L., & Vallortigara, G. (2010). Innate sensitivity for self-propelled causal agency in newly hatched chicks. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(9), 4483–4485. https://doi.org/10.1073/pnas.0908792107

Mascalzoni, E., Regolin, L., Vallortigara, G., & Simion, F. (2013). The cradle of causal reasoning: Newborns' preference for physical causality. *Developmental Science*, *16*(3), 327–335. https://doi.org/10.1111/desc.12018

Matsuno, T., & Tomonaga, M. (2017). Causal capture effects in chimpanzees (*Pan troglodytes*). *Cognition*, *158*, 153–164. https://doi.org/10.1016/j.cognition.2016.10.023

Meewis, F., Barezzi, I., Fagot, J., Claidière, N., & Dautriche, I. (2024). A comparative study of causal perception in Guinea baboons (*Papio papio*) and human adults. *PLOS ONE*, *19*(12), e0311294. https://doi.org/10.1371/journal.pone.0311294

Meewis, F., Fagot, J., Claidière, N., & Dautriche, I. (2025). Agent preference in chasing interactions in Guinea baboons (*Papio papio*): Uncovering the roots of subject–object order in language. *Psychological Science*, *36*(6), 465–477. https://doi.org/10.1177/09567976251344581

Meyerhoff, H. S., Schwan, S., & Huff, M. (2014). Perceptual animacy: Visual search for chasing objects among distractors. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(2), 702–717. https://doi.org/10.1037/a0034846

Michotte, A. (1963). *The perception of causality*. Routledge. https://doi.org/10.4324/9781315519050

Moors, P., Wagemans, J., & de-Wit, L. (2017). Causal events enter awareness faster than non-causal events. *PeerJ*, *5*, e2932. https://doi.org/10.7717/peerj.2932

Morey, R. D., Rouder, J. N., Jamil, T., Urbanek, S., Former, K., & Ly, A. (2015). *BayesFactor: Computation of Bayes Factors for Common Designs*. https://doi.org/10.32614/CRAN.package.BayesFactor

Muentener, P., & Bonawitz, E. (2017). The development of causal reasoning. In M. R. Waldmann (Ed.), *The Oxford handbook of causal reasoning*. Oxford University Press, 677–698. https://doi.org/10.1093/oxfordhb/9780199399550.013.40

Muentener, P., & Carey, S. (2010). Infants' causal representations of state change events. *Cognitive Psychology*, *61*(2), 63–86. https://doi.org/10.1016/j.cogpsych.2010.02.001

Napoli, D. J., & Sutton-Spence, R. (2014). Order of the major constituents in sign languages: Implications for all language. Frontiers in Psychology, *5*, 376. https://doi.org/10.3389/fpsyg.2014.00376

Oakes, L. M. (1994). Development of infants' use of continuity cues in their perception of causality. *Developmental Psychology*, *30*, 869–879. https://doi.org/10.1037/0012-1649.30.6.869

Oakes, L. M. (2010). Using habituation of looking time to assess mental processes in infancy. *Journal of Cognition and Development*, *11*(3), 255–268.

Oakes, L. M. (2017). Sample size, statistical power, and false conclusions in infant looking-time research. *Infancy*, *22*(4), 436–469. https://doi.org/10.1111/infa.12186

Ohl, S., & Rolfs, M. (2024). Visual routines for detecting causal interactions are tuned to motion direction. *eLife*, *13*, RP93454. https://doi.org/10.7554/eLife.93454.1

Oakes, L. M., Sperka, D., DeBolt, M. C., & Cantrell, L. M. (2019). Habit2: A stand-alone software solution for presenting stimuli and recording infant looking times in order to study infant development. *Behavior research methods*, *51*(5), 1943–1952.

Papeo, L., Vettori, S., Serraille, E., Odin, C., Rostami, F., & Hochmann, J.-R. (2024). Abstract thematic roles in infants' representation of social events. *Current Biology*, *34*(18), 4294–4300. https://doi.org/10.1016/j.cub.2024.07.081

R Core Team. (2021). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rissman, L., & Majid, A. (2019). Thematic roles: Core knowledge or linguistic construct? *Psychonomic Bulletin and Review*, *26*(6), 1850–1869. https://doi.org/10.3758/s13423-019-01634-5

Rochat, P., Striano, T., & Morgan, R. (2004). Who is doing what to whom? Young infants' developing sense of social causality in animated displays. *Perception*, *33*(3), 355–369. https://doi.org/10.1068/p3389

Rolfs, M., Dambacher, M., & Cavanagh, P. (2013). Visual adaptation of the perception of causality. *Current Biology*, *23*(3), 250–254. https://doi.org/10.1016/j.cub.2012.12.017

Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences*, *4*(8), 299–309. https://doi.org/10.1016/S1364-6613(00)01506-0

Silvey, C., Dienes, Z., & Wonnacott, E. (2024). Bayes factors for logistic (mixed-effect) models. *Psychological Methods*, https://doi.org/10.1037/met0000714

Strickland, B. (2017). Language reflects "core" cognition: A new theory about the origin of cross-linguistic regularities. *Cognitive Science*, *41*(1), 70–101. https://doi.org/10.1111/cogs.12332

Ünal, E., Ji, Y., & Papafragou, A. (2021). From event representation to linguistic meaning. *Topics in Cognitive Science*, *13*(1), 224–242. https://doi.org/10.1111/tops.12475

Ünal, E., Wilson, F., Trueswell, J., & Papafragou, A. (2024). Asymmetries in encoding event roles: Evidence from language and cognition. *Cognition*, *250*, 105868. https://doi.org/10.1016/j.cognition.2024.105868

Vauclair, J., Fagot, J., & Hopkins, W. D. (1993). Rotation of mental images in baboons when the visual input is directed to the left cerebral hemisphere. *Psychological Science*, *4*(2), 99–103. https://doi.org/10.1111/j.1467-9280.1993.tb00468.x

Völter, C. J., & Huber, L. (2021). Dogs' looking times and pupil dilation response reveal expectations about contact causality. *Biology Letters*, *17*(12), 20210465. https://doi.org/10.1098/rsbl.2021.0465

Webb, A., Knott, A., & MacAskill, M. R. (2010). Eye movements during transitive action observation have sequential structure. *Acta Psychologica*, *133*(1), 51–56. https://doi.org/10.1016/j.actpsy.2009.09.001

Wilson, F., Papafragou, A., Bunger, A., & Trueswell, J. (2011). Rapid extraction of event participants in caused motion events. In *Proceedings of the Annual Meeting of the Cognitive Science Society*.

Wilson, V. A. D., Zuberbühler, K., & Bickel, B. (2022). The evolutionary origins of syntax: Event cognition in nonhuman primates. *Science Advances*, *8*(25), eabn8464. https://doi.org/10.1126/sciadv.abn8464

Young, M. E., Beckmann, J. S., & Wasserman, E. A. (2006). Pigeons's discrimination of Michotte's launching effect. *Journal of the Experimental Analysis of Behavior*, *86*(2), 223–237. https://doi.org/10.1901/jeab.2006.60-05

Zuberbühler, K. (2019). Evolutionary roads to syntax. *Animal Behaviour*, *151*, 259–265. https://doi.org/10.1016/j.anbehav.2019.03.006

Zuberbühler, K. (2020). Syntax and compositionality in animal communication. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1789), 20190062. https://doi.org/10.1098/rstb.2019.0062

Zuberbühler, K. (2022). Event parsing and the origins of grammar. *WIREs Cognitive Science*, *13*(3), e1587. https://doi.org/10.1002/wcs.1587

Zuberbühler, K., & Bickel, B. (2022). Transition to language: From agent perception to event representation. *WIREs Cognitive Science*, *13*(6), e1594. https://doi.org/10.1002/wcs.1594

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Supporting Information