



HAL
open science

Infants' understanding of the causal power of agents and tools

Parvaneh Adibpour, Jean-Rémy Hochmann

► **To cite this version:**

Parvaneh Adibpour, Jean-Rémy Hochmann. Infants' understanding of the causal power of agents and tools. *Proceedings of the National Academy of Sciences of the United States of America*, 2023, 120 (50), pp.e2309669120. 10.1073/pnas.2309669120 . hal-04383707

HAL Id: hal-04383707

<https://hal.science/hal-04383707v1>

Submitted on 9 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Infants' understanding of the causal power of agents and tools

Parvaneh Adibpour,^{1,2*} & Jean-Rémy Hochmann^{3,4*}

¹ Université Paris Cité, NeuroDiderot Unit UMR1141, INSERM, 75019, Paris, France

² Université Paris Saclay, CEA, NeuroSpin, UNIACT, F-91191 Gif-sur-Yvette, France

³ CNRS UMR5229 - Institut des Sciences Cognitives Marc Jeannerod, 67 Boulevard Pinel, 69675, Bron, France.

⁴ Université Lyon 1 Claude Bernard, 69100, Villeurbanne, France

**To whom correspondence may be addressed:* E-mail: hochmann@isc.cnrs.fr, parvaneh.adibpour@inserm.fr

Significance Statement

Tool-use understanding is classically thought to involve reasoning about the causal mechanisms that relate tools and their effects (e.g., moving food closer). Through cultural evolution and accumulated innovations, however, humans have developed tools that became causally opaque (e.g., touch screens). We show that cultural transmission remains possible because even young infants are equipped with the fundamental capacity to represent arbitrary causal events where an agent acts on the world with a tool. Infants attributed to a ball the power to cause an arbitrary state change but only if that ball was itself launched by an agent. Infants thus consider that the power to cause arbitrary events ultimately belongs to intentional agents but agents can transfer this power to tools.

Abstract

Tools are objects that are manipulated by agents with the intention to cause an effect in the world. We show that the cognitive capacity to understand tools is present in young infants, even if these tools produce arbitrary, causally opaque effects. In Experiments 1-2, we used pupillometry to show that 8-month-old infants infer an invisible causal contact to account for the –otherwise unexplained– motion of a ball. In Experiments 3, we probed 8-month-old infants' account of a state change event (flickering of a cube) that lies outside of the explanatory power of intuitive physics. Infants repeatedly watched an intentional agent launch a ball behind an occluder. After a short delay, a cube, positioned at the other end of the occluder began flickering. Rare unoccluded events served to probe infants' representation of what happened behind the occluder. Infants exhibited larger pupil dilation, signaling more surprise, when the ball stopped before touching the cube, than when it contacted the cube, suggesting that infants inferred that the cause of the state change was contact between the ball and the cube. This effect was canceled in Experiment 4, when an inanimate sphere replaced the intentional agent. Altogether, results suggest that, in the infants' eyes, a ball (an inanimate object) has the power to cause an arbitrary state change, but only if it inherits this power from an intentional agent. Eight-month-olds are thus capable of representing complex event structures, involving an intentional agent causing a change with a tool.

Introduction

Tools are objects that are manipulated by agents with the intention to cause an effect in the world. Early tools in primate and human evolution, as in nowadays wild animals, were used to acquire food, groom conspecifics, or for self-defense [1]. In these situations, the causal mechanisms that relate tool use and their effects (e.g., moving food closer, removing parasites) are largely transparent and can be accounted for by intuitive theories [2-6]. But modern humans show a unique flexibility and versatility in their use of tools. Through cultural evolution, each generation improved the previous generation's tools, so much that underlying causal mechanisms have become opaque. The spectrum of possible effects of modern tools is virtually infinite, from switching on the light, to playing music or having Granma appear on a computer screen. This opacity may constitute a learning problem for children, especially if one considers that tool use requires reasoning about the causal mechanisms relating the tool and its effects [7]. In the present study, we propose that cultural transmission of tools is made possible by infants' disregard of causal opacity and an early understanding of the large causal power of agents and tools.

We asked whether infants understand that an intentional agent can cause an arbitrary state change on a third object through the use of an intermediary object, the tool. This understanding involves the integration, in a complex structure, of a chain of causal relations and three thematic roles: the *agent* acts on the *tool* (or *instrument*), which in turn affects the *patient's* state.

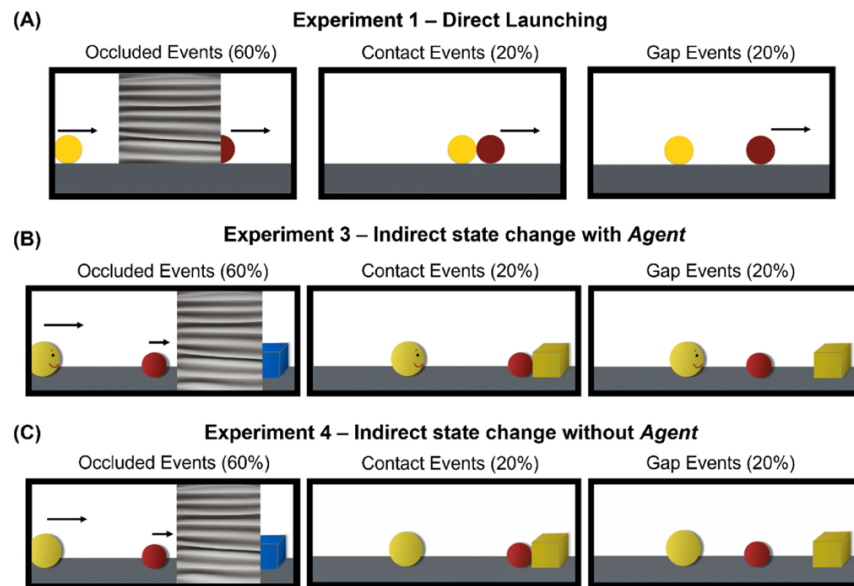


Fig. 1 – Schematic representation of the stimuli used in Experiments 1-4. In Experiments 1, 3 and 4, infants saw 60% of occluded events, 20% of contact events and 20% of gap events. Experiment 1 (A) tested causal inference for direct launching events. In Experiment 2 (not depicted), infants only saw the contact (50%) and gap (50%) events from Experiment 1. Experiments 3 (B) and 4 (C) tested causal inference for indirect state change events. The yellow object had face features and was introduced as a self-propelled intentional agent in Experiment 3 but not in Experiment 4.

Before addressing the issue of tool use, Experiment 1 introduced a method relying on pupillometry to replicate classic findings showing that infants infer a causal interaction, the contact between two objects, to account for the visible motion of a ball [8-11]. The inference of contact and temporal contiguity is taken as a signature of understanding the causal relationship between two events [12-14]. In contrast to classic looking time studies, pupillometry provides timing information, which can improve the interpretation of the results [15]. Eight-month-old infants (N = 20) watched repeated movies where, shortly after a yellow ball went behind an occluder, from the left, a second ball initially positioned at the right boundary of the occluder began to move (Fig. 1A). Rare interleaved unoccluded events served to probe infants' representation of what happened behind the occluder. Previous research showed that unexpected events trigger pupil dilation in infants, indicating surprise [15-16]. If infants inferred a contact between the two balls behind the occluder, they should be more surprised – exhibiting larger pupil dilation – by rare unoccluded Gap movies, in which the first ball did not contact the second ball, than by rare unoccluded Contact movies, in which the first ball contacted the second. To confirm that responses to unoccluded movies related to the interpretation of occluded movies, and not to the mere visual properties of Gap and Contact movies, a second group of infants (Experiment 2, N = 21) only watched the Contact and Gap movies. If Gap movies elicit more pupil dilation than Contact movies, independently from the processing of occluded movies, the same results should be obtained in Experiment 2 as in Experiment 1. Instead, we predicted larger pupil dilation for Gap compared to Contact events in Experiment 1 than in Experiment 2.

Next, in Experiment 3, we studied infants' interpretation of a tool-use event, where an agent used a ball to trigger the flickering of a cube. A prior study concluded that infant cognition does not encompass the notion that a mere contact between two inanimate objects could instigate a state change such as flickering [11]. This finding suggests that state-changes lie beyond the explanatory power of intuitive physics, which seems primarily focused on elucidating object motions. In contrast, infants consider that an intentional agent – e.g. a human hand – can elicit a state-change upon direct contact [11,17]. Here, we tested whether infants consider that agents can elicit state-change with a tool. Eight-month-old infants (N = 40) were first introduced with a self-propelled intentional agent [18]. Next, they watched occluded movies in which the agent launched a ball, which went behind an occluder, from the left (Fig. 1B). On the other end of the occluder, a partly occluded cube began flickering after a short delay. Again, interleaved unoccluded events served to probe infants' representation of what happened behind the occluder. In rare unoccluded Contact movies, infants watched the agent launching the ball that then contacted the cube, which immediately began flickering. In rare unoccluded Gap movies, infants watched the agent launching the ball, which stopped *before* contacting the cube. The cube began flickering with a delay. If infants conceived that the ball is the proximate cause of the cube's flickering, they should have inferred a contact between the ball

and the cube behind the occluder, and temporal contiguity between that contact and flickering. In consequence, Gap movies should be more surprising and elicit larger pupil dilation than Contact movies. Finally, tools are defined as objects that are manipulated by agents with the intention to cause an effect. Following this definition, we predict that the representation of a tool-use event necessitates the involvement of an intentional agent. In Experiment 4, infants (N = 40) watched the same sequence of movies as in Experiment 3, except that a yellow sphere, with no indication of agency, replaced the intentional agent (Fig., 1C). We predicted weaker or null effects in this group.

Results

Pupil dilation results are displayed in Fig. 2.

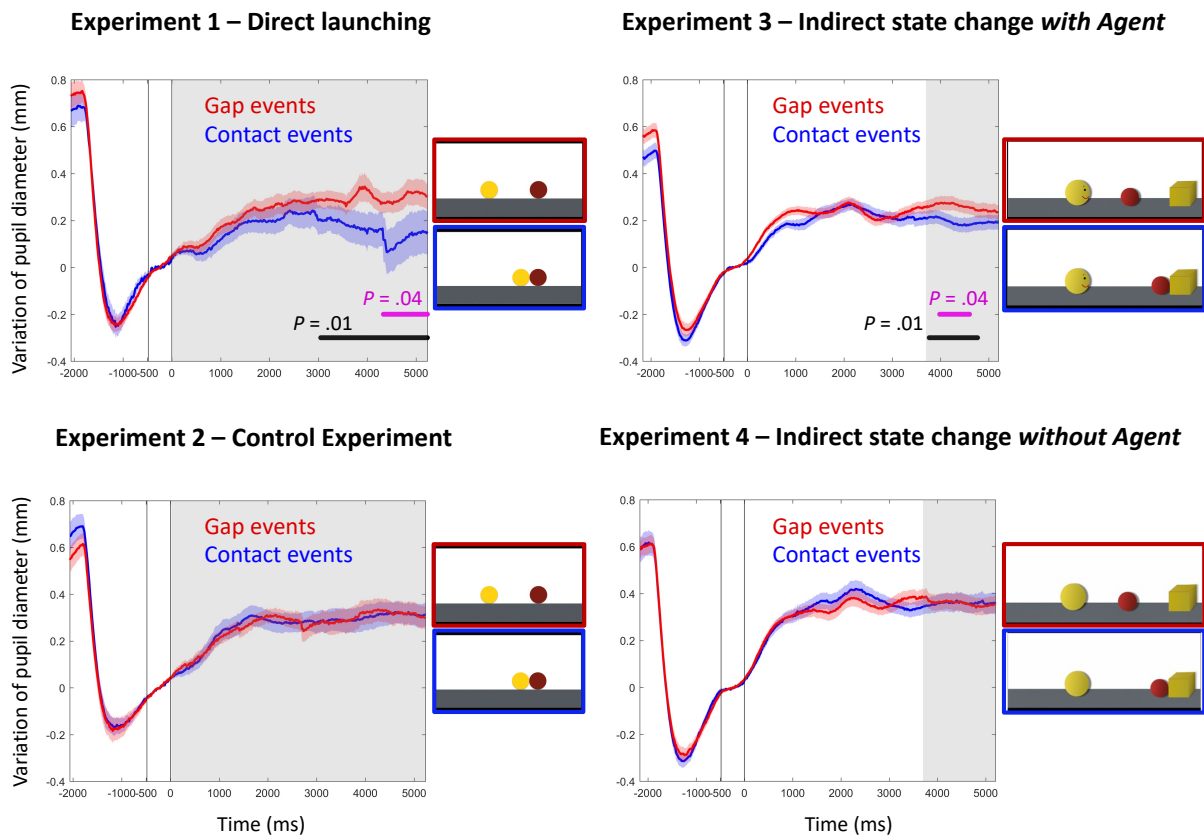


Fig. 2 –Pupil dilation for the Contact (blue) and Gap (red) events in Experiments 1-4. The baseline is taken in the last 500 ms before the two types of events differ (i.e., when the ball stops in Gap events). Shaded gray areas indicate the time window of interest in which cluster mass permutation tests were performed. For Experiments 1 and 2, this time window extended from the end of the baseline (0 ms) to the end of the trial. For Experiments 3 and 4, the time window of interest was restricted to the period where significant effects were observed in Experiments 1-2. Horizontal black bars indicate periods of significant larger pupil dilation for Gap vs Contact events in each Experimental Group, analyzed individually. Horizontal magenta bars indicate periods of significant interaction between Experimental Group and Trial Type. Light color areas indicate standard errors from the mean.

Experiments 1-2 – In Experiment 1, 8-month-old infants (N=20) watched repeated launching events involving two balls. Contact events (20%), when the first ball contacted the second, and Gap events (20%), when the first ball stopped before touching the second, were intermixed with a majority of Occluded events (60%) that were compatible with both the Contact and Gap events, as the fate of the first ball was hidden by an occluder (Fig. 1). Infants exhibited larger pupil dilation for Gap events than for Contact events between 3050 and 5233 ms (cluster mass permutation test, $P = .01$ one-tail). In Experiment 2, 8-month-old infants (N=21) only saw unoccluded Gap (50%) and Contact (50%) events. There was no significant difference of pupil dilation in response to the two types of events.

Analyzing Experiments 1 and 2 together, infants overall exhibited larger pupil dilation for Gap than for Contact events between 3700 and 5200 ms (cluster mass permutation test, $P = .02$ one-tail). In addition, the interaction between Experimental Group (Experiment 1, Experiment 2) and Trial Type (Gap, Contact) was significant between 4333 and 5233 ms (cluster mass permutation test, $P = .04$ one-tail); i.e., the difference of pupil dilation between Gap and Contact events was larger in Experiment 1 than in Experiment 2. In fact, in that time window, pupil dilation for Gap relative to Contact events was significant in Experiment 1 ($t(19) = 2.46$; $P = .012$ one-tail) but not in Experiment 2 ($t(20) = .06$, $P = .47$ one-tail).

Experiments 3-4 – In Experiment 3, 8-month-old infants (N=40) watched repeated state change events, involving an intentional agent, a ball and a flickering cube. All events began with the agent launching the ball. Unoccluded Contact events (20%), when the ball contacted the cube, and Gap events (20%), when the ball stopped before contacting the cube, were intermixed with a majority of Occluded events (60%) that were compatible with both the Contact and Gap events, as the fate of the ball was hidden by an occluder (Fig. 1). The analysis was restricted to the time window of interest defined by the effects observed in Experiments 1-2 between 3700 after the gap onset and the end of the trial. Infants exhibited larger pupil dilation for Gap events than for Contact events between 3767 and 4766 ms (cluster mass permutation test, $P = .01$ one-tail). In Experiment 4 (N=40), instead of an intentional agent, an inanimate yellow sphere launched the ball. There was no significant difference of pupil dilation in response to Contact and Gap events.

Analyzing Experiments 3 and 4 together, infants overall exhibited larger pupil dilation for Gap than for Contact events between 3983 and 4600 ms (cluster mass permutation test, $P = .03$ one-tail). In addition, the interaction between Experimental Group (Experiment 3, Experiment 4) and Trial Type (Gap, Contact) was significant between 4083 and 4583 (cluster mass permutation test, $P = .04$ one-tail); i.e. the difference of pupil dilation between Gap and Contact events was larger in Experiment 3 than in Experiment 4. In fact, in that time window, pupil dilation for Gap relative to Contact events was significant in Experiment 3 ($t(39) = 2.8$; $P = .004$ one-tail) but not in Experiment 4 ($t(39) = .04$; $P = .48$ one-tail).

Trial Number analysis – The pupil dilation effects reported in the above analyses are thought to be equivalent to longer looking times in traditional violation of expectation paradigms. In the current paradigm, trials had a fixed duration, making cumulative looking time analyses uninformative. Nevertheless, we explored the equivalence between pupil dilation and increased looking time by analyzing differences in the number of Gap and Contact trials included in the final analysis of the pupil data. Indeed, each trial was only included if the infant looked long enough at the screen (i.e., more than 65% of the time window of interest, see *Analysis* section). A higher number of included Gap trials would thus suggest that infants tended to look longer in Gap trials than in Contact trials. Table 1 summarizes the number of Gap and Contact trials infants contributed to the final analyses. For each participant, we computed a difference score ($\# \text{good Gap trials} - \# \text{good Contact trials} / (\# \text{good Gap trials} + \# \text{good Contact trials})$); a positive score indicated more Gap trials; a negative score indicated more Contact trials. We then compared the distribution of this score to chance (0). This analysis showed that infants contributed significantly more Gap than Contact trials in Experiment 1 ($t(19) = 2.6$; $P = .009$) and in Experiment 3 ($t(39) = 2.7$; $P = .006$), while there was no difference in Experiment 2 ($t(20) = .1$; $P = .46$) and in Experiment 4 ($t(39) = .3$; $P = .36$). The effects in Experiments 1 and 3 cannot result from a trial order effect, as Gap and Contact trials were intermixed. Rather, we propose that these effects reflect a tendency to look longer at Gap trials. In keeping with the interpretation of traditional violation of expectation paradigms, longer looking times indicate surprise. The trial number analysis thus converges with the pupillometry analysis, suggesting that infants are surprised by Gap trials in Experiments 1 and 3, but not in Experiments 2 and 4.

Table 1 - Average number of Contact and Gap trials that infants saw, and average number of good trials (see *Analysis*), included in the final analyses.

Experiment	#Contact trials	#good Contact trials	#Gap trials	#good Gap trials
Experiment 1 (Direct launching)	8.4 (2.6)	2.9 (2.1)	8.5 (2.4)	3.8 (2.5)
Experiment 2 (Control)	9.5 (1.3)	5.2 (2.8)	9.4 (1.5)	5.3 (2.8)
Experiment 3 (Indirect state change with Agent)	8.5 (1.9)	5.4 (2.6)	8.8 (1.8)	5.8 (2.1)
Experiment 4 (Indirect state change without Agent)	7.6 (2.3)	5.0 (2.3)	7.7 (2.2)	5.0 (2.6)

Standard deviations from the mean are indicated in parentheses.

Discussion

The present studies explored how infants conceive the causal power of agents, objects and tools. Experiments 1-2 showed that 8-month-old infants spontaneously attribute to a ball the power to cause another object’s motion through contact, respecting the laws of intuitive physics. In addition, these

experiments validated our pupillometry methods. Next, Experiments 3-4 showed that 8-month-old infants attribute to a ball the power of causing an arbitrary (not explained by any obvious laws of physics) state change to a cube, but only if the ball was itself launched by an intentional agent. In other words, the inference of the final causal interaction between the ball and the cube depended on the causal launching between the agent and the ball. This fact suggests that the two causal interactions were not represented as two independent sequential events, but were rather integrated into one complex structure: the agent caused the final state change *with* the ball. The ultimate cause of the state change is thus the agent, who transmitted their causal power to the ball.

The distinction between inanimate objects and intentional agents is a core structure of the mind [3,19]. Infants expect inanimate objects to be moved by causal external forces [9-10, 20], while intentional agents exert self-propelled motion and rational actions [21] guided by internal goals [22], create order [23-26] and are prone to social affiliation [27-31] and moral evaluation [32-34]. Our results further support an additional fundamental conceptual distinction between intentional agents and objects: their respective causal powers (see also [11, 17-18]). While the causal power of objects is limited to intuitive physics and possibly to direct actions (Experiments 1 and 4), the power of intentional agents extends to arbitrary state changes and indirect instrumental actions (Experiment 3). In addition, we show that an inanimate object can inherit the agent's causal power, when it acquires the role of a tool, mediating between an agent and a final effect. This suggests that, conceptually, tools occupy a zone of contact between the core domains of inanimate objects and that of intentional agents.

However, our results should not be interpreted as meaning that the causal power of agents and tools is fully unconstrained. In particular, the results reported here and elsewhere [11] suggest that young infants still expect agents to act by contact, whether directly or through a tool, and the effect to be immediate (temporal contiguity). The primacy of contact and temporal contiguity suggests that the schematic representation of causal interactions involving agents may be derived from those of intuitive physics, maintaining some of their constraints, while opening up the spectrum of possible effects (not just motion). How (and when) infants come to understand contactless tools, such as the TV remote control, is a question for future studies. Note however, that in our experiments, contact and temporal contiguity covaried. It remains possible that infants would accept contactless causal actions in presence of temporal contiguity, and if a goal is clearly apparent [35]. More generally, future work is needed to describe the spectrum of agents' possible actions and circumscribe their causal power.

A long tradition in developmental psychology has argued that understanding instrumental actions, including those involving a tool, develops towards the end of the first year of life and consolidates in the second year of life, in parallel with children's first-person motor behavior (e.g., [36-38]). Those results have been used to support views of action understanding as relying on motor simulation [39-40]. In contrast,

converging with observations that object function supports early object individuation [41-42], our results show that 8-month-old infants already possess the cognitive capacities to understand instrumental actions. Far from constituting a learning problem, the opacity of the causal mechanism that relates contact with the state change may have allowed infants to by-pass simulation mechanisms and demonstrate their early understanding of tools. In sum, the notion of a tool appears to be a basic element of the mind, present early in infancy, prior to children's production of instrumental actions. Incidentally, our results further suggest that the representation of a thematic role *instrument* precedes the acquisition of explicit morphosyntactic markers of that role [43-44].

Humans are in effect the only species to have developed complex causally-opaque tools. Even so, we cannot exclude that the capacities observed here in young infants are shared with other species, especially non-human primates. Wild chimps, for instance, have demonstrated an understanding of cooking devices [45], and laboratory monkeys are capable to acquire modern tools such as joysticks and touchscreens, albeit after extended trainings periods [46]. Yet, subtle differences may exist between species regarding their resilience to causal opacity, and their need for a detailed understanding of the causal mechanisms involved in tool use. The phenomenon known as over-imitation – i.e., the assumption that all elements in an observed action sequence play a causal (even if opaque) role in fulfilling the artifacts' function – was proposed as a mechanism that allows children to learn the use of novel tools, even when causal mechanisms are not readily apparent [47-50]. Empirical studies found that over-imitation is weaker in non-human primates than in 3- and 5-year-old human children, suggesting reduced resilience to causal opacity [51-52], hence possibly higher difficulty in representing arbitrary causal events. Further comparative studies are needed to progress on these issues.

Finally, our results highlight the potential interest of the timing information provided by pupillometry in investigating infants' reasoning. It holds the promise to lift (in part) the ambiguity held in the interpretation of classical looking-time measures in the violation of expectation paradigm. The name itself of that paradigm implies that infants *expect* a precise event and are surprised when something else happens; here, they would predict a contact between the ball and the cube, and be surprised when seeing a gap between the two objects. Increased looking time would thus signal the detection of a prediction error. Nevertheless, the same effects are also generally compatible with *a posteriori* evaluations of observed events, in which infants do not predict what will happen but only detect incongruencies between events *a posteriori*. In Experiment 3, the incongruency could be the observation of flickering without an apparent cause in gap events. When pupil dilation is triggered by perceptual changes and/or the violations of well-defined predictions, the effects are registered around 1000-1500 ms in infants [15-16,53-54]. In comparison, the pupil responses observed here occurred late – i.e., starting more than 3000 or 4000 ms after the ball stopped in Gap events, that is about 2000 ms later than what simple perceptual predictions would elicit.

Such late effects make it unlikely that infants reacted to the gap because they formed a detailed prediction of a contact between the ball and the cube. A possible scenario is that the state-change in occluded events triggered curiosity, asking “why is the cube flickering?”, and unoccluded events provided alternative answers that infants were capable of evaluating. In this scenario, pupil dilation shows that infants considered a contact between the ball and the cube a better explanation for the flickering than a spatiotemporal gap, provided the ball was launched by an agent. The inference of a contact as the cause of flickering may result from these *a posteriori* evaluations.

In conclusion, along cultural evolution, humans have produced tools, whose power transcends the limits of intuitive physics, to the point of causing apparently arbitrary effects. Our results suggest that the abilities to create and understand these tools are rooted in an early capacity to understand and represent even arbitrary and opaque causal effects, as long as an intentional agent can be conceived as the ultimate cause.

Material and Methods

Participants: Sample size in Experiment 1 followed a power analysis of the results of [11] (Experiment 1) conducted on G*Power. That experiment replicated a previous study [8], with a sample size of 20 infants. An effect size of $d = 1.11$ was computed from a t-value of 2.42 and degree of freedom of 19. To obtain a similar effect a sample size of $N > 11$ is necessary. We thus kept the sample size of $N = 20$ participants for Experiments 1 and 2.

Previous studies reported no significant effect when an inanimate object caused a state change [11]. Moreover, an instrumental action is intuitively more complex than a direct action, as it involves three objects instead of two, and two relations instead of one. Therefore, we considered that the task in Experiment 3 could be more difficult than in Experiment 1, possibly yielding a smaller effect. We thus doubled the sample size to $N = 40$ participants for Experiments 3 and 4.

Experiment 1 - Twenty infants participated in Experiment 1 (*mean age* = 8 months 11 days, range: 7 months 29 days to 8 months 22 days). Four additional infants were tested but excluded from the final analysis for not providing good data for at least one Contact movie and one Gap movie. One additional infant was excluded because of parental interference with the procedure.

Experiment 2 - Twenty-one infants participated in Experiment 2 (*mean age* = 8 months 11 days, range: 8 months 1 day to 8 months 28 days). One additional infant was tested but excluded from the final analysis for not providing good data for at least one Contact movie and one Gap movie.

Experiment 3 - Forty infants participated in Experiment 3 (*mean age* = 8 month 12 days, range: 8 months 0 day to 9 months 0 day). Two additional infants were tested but excluded from the final analysis for not providing good data for at least one Contact movie and one Gap movie.

Experiment 4 - Forty infants participated in Experiment 4 (*mean age* = 8 months 13 days, range: 8 months 1 day to 9 months 0 day). Two additional infants were tested but excluded from the final analysis for not providing good data for at least one Contact movie and one Gap movie.

All participants were recruited through the consultation of birth records at the local city halls. This study was approved by the local ethics committee “Comité de Protection des Personnes Sud Est II”. Parents gave an informed consent before participating in the study and were informed that they could withdraw their consent at any time without any consequence.

Stimuli: Short movies for Experiments 1-4 were created using Keynote software. Stimuli in Experiments 1-2 are adapted from [11]. Three movies were created. At the beginning of the Occluded movie, a red ball was visible on the screen, positioned at the right end of and partly occluded by an occluder. A yellow ball, initially placed on the left of the screen began moving at 1600 ms and went behind the occluder at 2030 ms. At 2400 ms, the red ball began moving rightward and left the screen at 3060 ms. The movie ended at 4400 ms. The Contact and Gap movies showed the same sequence of events, except that the occluder was removed, and each movie showed a different scenario of what happened behind the occluder. In the Contact movie, the yellow ball stopped when contacting the red ball at 2400 ms. The red ball began moving immediately. In the Gap movie, the yellow ball stopped short of touching the red ball at 2100 ms. The red ball began moving at 2400 ms. Contact and Gaps movies ended at 7333 ms

Four movies were created for Experiment 3. In the Introduction movie, a yellow agent taking the shape of a sphere with two eyes and a mouth entered the screen from the left. It paused at 1700 ms, turned to face the participant and jumped. Then it moved right ward. It repeated the same actions again at 4330 and 6930 ms. The Introduction movie lasted 11860 ms. At the beginning of the Occluded movie, a red ball was placed at the left of an occluder, which partly occluded a blue cube on its right end. The yellow agent entered the screen from the left. It stopped at 1760 ms, when contacting the red ball, which began moving immediately. The red ball went behind the occluder at 1800 ms. The cube began alternating between blue and yellow at 2530 ms until 3330 ms. This color alternation was accompanied by a repeated glass sound. The Contact and Gap movies showed the same sequence of events, except that the occluder was removed, and each movie showed a different scenario of what happened behind the occluder. In the Contact movie, the red ball stopped when contacting the cube at 2530 ms. The cube began changing color immediately until 3330. In the Gap movie, the red ball stopped short of touching the cube at 2200 ms. The cube began changing color at 2530 ms until 3330. Contact and Gap movies ended at 7383 ms.

Three movies were created for Experiment 4. They had the exact same structure as respectively the Occluded, Contact and Gap movies of Experiment 3, except that the yellow agent was replaced by a yellow sphere that had no face feature.

Procedure: Presentation of the stimuli as well as data recordings were controlled by PsyScope X (<http://psy.cns.sissa.it>). Each trial in the experiment consisted in the presentation of one movie. Two trials were separated by an animation of colorful rotating arrows to attract the attention of participants toward the center of the screen. Each trial began when participants gazed toward the center of the screen for at least 100 ms.

Experiment 1 consisted in a maximum of 52 trials. Movies were shown in a pseudo-random order, with 60% of Occluded movies, 20% of Contact movies and 20% of Gap movies. The first 5 trials always showed Occluded movies; two Gap or two Contact movies never directly followed each other.

Experiment 2 consisted in 10 contact and 10 Gap movies showed in a pseudo-random order, avoiding more than two repetitions of the same movie.

Experiment 3 began with the Introductory movie, followed by up to 52 trials. Movies were shown in a pseudo-random order, with 60% of Occluded movies, 20% of Contact movies and 20% of Gap movies. The first 5 trials always showed Occluded movies; two Gap or two Contact movies never directly followed each other.

Experiment 4 consisted in a maximum of 52 trials. Movies were shown in a pseudo-random order, with 60% of Occluded movies, 20% of Contact movies and 20% of Gap movies. The first 5 trials always showed Occluded movies; two Gap or two Contact movies never directly followed each other.

Analysis:

All data and codes for analyses are shared publicly at osf.io/4caht/.

The pupil diameter for both eyes was recorded for gazes on the screen. For each trial, we considered a baseline time window lasting 500 ms and ending at the moment the Contact and Gap movies began to differ, that is when the ball stopped in the Gap movies (2100 ms after the trial onset in Experiments 1-2; 2200 ms after the trial onset in Experiments 3-4). The end of the baseline was then considered time 0.

We excluded trials with less than 65%[#] of pupil diameter information over the period that begins with the gap onset in Gap events and ends with the trial – i.e., 0-5233 ms in Experiments 1-2, 0-5183 ms in Experiments 3-4 – and/or less than 100 ms of pupil diameter information in the baseline time window. Only the Contact and Gap trials were analyzed. Infants with at least one included trial of each type (Contact and Gap) were included in the final analyses. The average number of good trials per infant is reported in Table 1 for each experiment. Missing data points for good trials were linearly interpolated.

Non-parametrical cluster mass permutation tests [53,55] were implemented to probe the variation of pupil dilation in response to Contact and Gap movies and the effect of Experimental Group (Experiment 1 vs.

[#] The choice of this parameter was set for the analysis of Experiment 1 and maintained for the other experiments. In *Supplementary Materials*, we report a multiverse analysis, demonstrating the solidity of our results when varying this criterion between 50% and 80%.

Experiment 2; Experiments 3 vs. Experiment 4). To test the effect of Trial Type (Gap, Contact), we compared the pupil dilation for Gap vs Contact trials at each timepoint with a paired-sample t-test. Neighboring timepoints showing an effect with a t-value larger than a threshold corresponding to ($p < 0.05$) were clustered and tested for significance using a permutation test, with 1000 random permutations of the trial-type labels on the original data. The significance probability of the original clusters was computed as the number of times the shuffled data produced clusters with higher summed t values than the real data. To test the interaction of Experimental Group (Experiment 1 vs. Experiment 2; Experiment 3 vs. Experiment 4) and Trial Type (Gap vs. Contact), we computed the time-course of the difference of pupil dilation between Gap and Contact trials for each infant. We then compared these values at each timepoint using 2-sample t-tests. Neighboring timepoints showing an effect with a t-value larger than a threshold corresponding to ($p < 0.05$) were clustered and tested for significance using a permutation test, with 1000 random permutations of the Experimental Group labels on the original data. The significance probability of the original clusters was computed as the number of times the shuffled data produced clusters with higher summed t values than the real data. The time window where effects were observed in Experiments 1 and 2 served to highlight a window of interest to constrain the analyses in Experiments 3 and 4.

All t-tests reported in the paper are one-tail, as we had strong predictions for each of the tests that we ran. We predicted that Gap trials should elicit larger pupil dilation than Contact trials in Experiments 1 and 3; we predicted larger differences between Gap and Contact trials in Experiment 1 than in Experiment 2, because we hypothesized that the effect observed in Experiment 1 is due to the interpretation of occluded events that are absent in Experiment 2; we predicted larger differences between Gap and Contact trials in Experiment 3 than in Experiment 4, because we hypothesized that infants consider that the ultimate cause of the state change in Experiment 3 is the Agent, which is absent in Experiment 4.

We also compared the number of Gap vs Contact trials included in each Experiment. This analysis is detailed in *Trial Number analysis* in the result section.

Acknowledgements

This work was funded by the collaborative McDonnell Foundation Grant 220020449 and the LABEX CORTEX (ANR-11-LABX-0042). We thank Liuba Papeo, Susan Carey and two anonymous reviewers for comments on previous versions of this manuscript. We thank Céline Spriet, Marion Gasselin, and Emilie Serraille for their help in recruiting and testing infants.

Author contributions: P.A. and J-R.H. designed research; performed research; analyzed data; and wrote the paper.

Competing Interests: The authors declare no competing interest.

References

1. R. W. Schumaker, K. R. Walkup, B. B. Beck (2011). *Animal Tool Behavior. The Use and Manufacture of Tools by Animals*, Johns Hopkins University Press
2. Wellman, H. M., & Gelman, S. A. (1992). Cognitive development: Foundational theories of core domains. *Annual Review of Psychology*, 43 (1), 337–375.
3. Carey, S. (2009). *The Origin of Concepts*. Oxford University Press
4. Spelke, E. S. (2022). *What Babies Know: Core Knowledge and Composition Volume 1* (Vol. 1). Oxford University Press.
5. Spelke, E. S. (1990). Principles of object perception. *Cognitive science*, 14(1), 29-56.
6. Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental science*, 10(1), 89-96.
7. Osiurak, F., & Reynaud, E. (2020). The elephant in the room: What matters cognitively in cumulative technological culture. *Behavioral and Brain Sciences*, 43, e156.
8. Ball, W. (1973). The perception of causality in the infant. Presented at the Meeting of the Society for Research in Child Development, Philadelphia, PA.
9. Kotovsky, L., & Baillargeon, R. (2000). Reasoning about collisions involving inert objects in 7.5-month-old infants. *Developmental science*, 3(3), 344-359.
10. Wang, S. H., Kaufman, L., & Baillargeon, R. (2003). Should all stationary objects move when hit? Developments in infants' causal and statistical expectations about collision events. *Infant Behavior and Development*, 26(4), 529-567.
11. Muentener, P., & Carey, S. (2010). Infants' causal representations of state change events. *Cognitive psychology*, 61(2), 63-86.
12. Leslie, A. M. (1982). The perception of causality in infants. *Perception*, 11(2), 173-186.
13. Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality?. *Cognition*, 25(3), 265-288.
14. Michotte, A. E. (1963). *The perception of causality* (TR Miles, Trans.). London, England: Methuen & Co.
15. Hochmann, J. R., & Toro, J. M. (2021). Negative mental representations in infancy. *Cognition*, 213, 104599.
16. Hochmann, J. R. (2022). Representations of Abstract Relations in Infancy. *Open Mind*, 1-20.
17. Liu, S., Brooks, N. B., & Spelke, E. S. (2019). Origins of the concepts cause, cost, and goal in prereaching infants. *Proceedings of the National Academy of Sciences*, 116(36), 17747-17752.
18. Saxe, R., Tenenbaum, J. B., & Carey, S. (2005). Secret agents: Inferences about hidden causes by 10- and 12-month-old infants. *Psychological Science*, 16(12), 995-1001.

19. Spriet, C., Abassi, E., Hochmann, J. R., & Papeo, L. (2022). Visual object categorization in infancy. *Proceedings of the National Academy of Sciences*, *119*(8), e2105866119.
20. Mascialoni, E., Regolin, L., & Vallortigara, G. (2010). Innate sensitivity for self-propelled causal agency in newly hatched chicks. *Proc. Natl. Acad. Sci. U. S. A.*, *107* (9), 4483 :4485.
21. Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naive theory of rational action. *Trends in cognitive sciences*, *7*(7), 287-292.
22. Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*(1), 1-34.
23. Kim, M., & Schachner, A. (2021, July). From music to animacy: Causal reasoning links animate agents with musical sounds. In *Proceedings of the Annual Conference of the Cognitive Science Society*.
24. Ma, L., & Xu, F. (2013). Preverbal infants infer intentional agents from the perception of regularity. *Developmental psychology*, *49*(7), 1330.
25. Newman, G. E., Keil, F. C., Kuhlmeier, V. A., & Wynn, K. (2010). Early understandings of the link between agents and order. *Proceedings of the National Academy of Sciences*, *107*(40), 17140-17145.
26. Schachner, A. (2017). From music to movement: The role of causal reasoning in infants' and adults' representations of music. Talk presented at the Society for Music Perception and Cognition, San Diego, CA
27. Powell, L. J., & Spelke, E. S. (2013). Preverbal infants expect members of social groups to act alike. *Proceedings of the National Academy of Sciences*, *110*(41), E3965-E3972.
28. Powell, L. J., & Spelke, E. S. (2018). Human infants' understanding of social imitation: Inferences of affiliation from third party observations. *Cognition*, *170*, 31-48.
29. Powell, L. J., & Spelke, E. S. (2018). Third-party preferences for imitators in preverbal infants. *Open Mind*, *2*(2), 61-71.
30. Bian, L., & Baillargeon, R. (2022). When Are Similar Individuals a Group? Early Reasoning About Similarity and In-Group Support. *Psychological Science*, *33*(5), 752-764.
31. Thomas, A. J., Woo, B., Nettle, D., Spelke, E., & Saxe, R. (2022). Early concepts of intimacy: Young humans use saliva sharing to infer close relationships. *Science*, *375*(6578), 311-315.
32. Sloane, S., Baillargeon, R., & Premack, D. (2012). Do infants have a sense of fairness?. *Psychological science*, *23*(2), 196-204.
33. Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, *450*(7169), 557-559.
34. Ting, F., Dawkins, M. B., Stavans, M., & Baillargeon, R. (2019). Principles and concepts in early moral cognition. *The social brain: A developmental perspective*, 41-65.
35. Southgate, V., Johnson, M. H., & Csibra, G. (2008). Infants attribute goals even to biomechanically impossible actions. *Cognition*, *107*(3), 1059-1069.
36. Piaget, J. (1953). *The origins of intelligence in the child*. London: Routledge & Kegan Paul.

37. Sommerville, J.A., Hildebrand, E.A., & Crane, C.C. (2008). Experience matters: The impact of doing versus watching on infants' subsequent perception of tool use events. *Developmental Psychology*, 44, 1249–1256
38. 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.
39. Hunnius, S., & Bekkering, H. (2014). What are you doing? How active and observational experience shape infants' action understanding. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1644), 20130490.
40. Woodward, A. L., & Gerson, S. A. (2014). Mirroring and the development of action understanding. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1644), 20130181.
41. Futó, J., Téglás, E., Csibra, G., & Gergely, G. (2010). Communicative function demonstration induces kind-based artifact representation in preverbal infants. *Cognition*, 117(1), 1-8.
42. Stavans, M., & Baillargeon, R. (2018). Four-month-old infants individuate and track simple tools following functional demonstrations. *Developmental Science*, 21(1), e12500.
43. Lidz, J., White, A. S., & Baier, R. (2017). The role of incremental parsing in syntactically conditioned word learning. *Cognitive Psychology*, 97, 62-78.
44. Rissman, L., & Majid, A. (2019). Thematic roles: Core knowledge or linguistic construct?. *Psychonomic bulletin & review*, 26(6), 1850-1869.
45. Warneken, F., & Rosati, A. G. (2015). Cognitive capacities for cooking in chimpanzees. *Proceedings of the Royal Society B: Biological Sciences*, 282(1809), 20150229.
46. Baraduc, P., Duhamel, J. R., & Wirth, S. (2019). Schema cells in the macaque hippocampus. *Science*, 363(6427), 635-639.
47. Hernik, M., & Csibra, G. (2009). Functional understanding facilitates learning about tools in human children. *Current opinion in neurobiology*, 19(1), 34-38.
48. Lyons, D. E., Young, A. G., & Keil, F. C. (2007). The hidden structure of overimitation. *Proceedings of the National Academy of Sciences*, 104(50), 19751-19756.
49. Lyons, D. E., Damrosch, D. H., Lin, J. K., Macris, D. M., & Keil, F. C. (2011). The scope and limits of overimitation in the transmission of artefact culture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1567), 1158-1167.
50. Clay, Z., & Tennie, C. (2018). Is overimitation a uniquely human phenomenon? Insights from human children as compared to bonobos. *Child development*, 89(5), 1535-1544.
51. Horner, V., & Whiten, A. (2005). Causal knowledge and imitation/emulation switching in chimpanzees (Pan troglodytes) and children (Homo sapiens). *Animal cognition*, 8, 164-181.

52. McGuigan N., Whiten A., Flynn E., Horner V. (2007). Imitation of causally-opaque versus causally-transparent tool use by 3- and 5-year-old children. *Cogn. Dev.* **22**, 353–364
53. Hochmann, J-R., & Papeo, L. (2014). The invariance problem in infancy: a pupillometry study. *Psychological Science* 25, 2038-2046.
54. Wetzel, N., Buttelmann, D., Schieler, A., & Widmann, A. (2016). Infant and adult pupil dilation in response to unexpected sounds. *Developmental psychobiology*, 58(3), 382-392.
55. Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods* 164, 177-190.

Supplementary materials

Multiverse Analysis

In order to assess the robustness of our results, we conducted a *multiverse analysis*, varying the arbitrary criterion for selecting good trials between 50% and 80%. While 50% is probably too lenient, including trials for which infants are not much attentive, 80% resulted in too high an attrition rate in Experiment 1 (about 50%). The chosen value of 65%, as reported in the main text, constitutes a trade-off between these two values for Experiment 1 (a conceptual replication of several looking-time studies [1-4]). It was set for the analysis Experiment 1 and maintained for the other experiments.

For each value of the exclusion criterion, for Experiments 1-4, Table S1 reports the results of a cluster mass permutation test assessing larger pupil dilation for Gap vs. Contact events, and Table S2 reports the results of the analysis of the number of contributed Gap and Contact trials. These analyses show that the results – especially those of Experiments 3-4 – are largely independent from the choice for the value of the exclusion criterion. This multiverse analysis thus supports the main conclusions presented in the main text.

Table S1 – Summary results of a *multiverse analysis* of the pupil dilation. The criterion for trial exclusion is a minimal percentage of pupil data required to include a trial in the final analysis. For each experiment and each value of the criterion for trial exclusion, we report the result of a cluster mass permutation test assessing larger pupil dilation in Gap vs. Contact events. The time window where the effect is observed is indicated when the P-value is significant ($P < .05$) or when it shows a trend ($P < .10$). n.s.: not significant; i.e., $P > .10$. * $P < .05$

	Criterion for trial exclusion						
	50%	55%	60%	65%	70%	75%	80%
Experiment 1	N=23 $P = \text{n.s.}$	N=21 $P = .07$ 4333-5000 ms	N=21 $P = .07$ 4333-5000 ms	N=20 $P = .01^*$ 3050-5233	N=17 $P = .08$ 3600-4233 ms	N=13 $P = .06$ 3566-4300 ms	N=12 $P = \text{n.s.}$
Experiment 2	N=22 $P = \text{n.s.}$	N=22 $P = \text{n.s.}$	N=21 $P = \text{n.s.}$	N=21 $P = \text{n.s.}$	N=20 $P = \text{n.s.}$	N=18 $P = \text{n.s.}$	N=16 $P = \text{n.s.}$
Experiment 3	N=40 $P = .01^*$ 3700-4916 ms	N=40 $P = .005^*$ 3700-4950 ms	N=40 $P = .01^*$ 3700-4700 ms	N=40 $P = .01^*$ 3766-4766 ms	N=40 $P = .008^*$ 3700-4966 ms	N=38 $P = .004^*$ 3766-5150 ms	N=36 $P = .003^*$ 3866-5150 ms
Experiment 4	N=41 $P = \text{n.s.}$	N=40 $P = \text{n.s.}$	N=40 $P = \text{n.s.}$	N=40 $P = \text{n.s.}$	N=39 $P = \text{n.s.}$	N=38 $P = \text{n.s.}$	N=36 $P = \text{n.s.}$

Table S2 – Summary results of a *multiverse analysis* of the number of trial for each Trial Type. The criterion for trial exclusion is a minimal percentage of pupil data required to include a trial in the final analysis. For each experiment and each value of the criterion for trial exclusion, we report the result of an one-sample t-test assessing whether infants contributed more trials in the Gap condition than in the Contact condition. * $P < .05$

	Criterion for trial exclusion						
	50%	55%	60%	65%	70%	75%	80%
Experiment 1	N=23 $P = .24$	N=21 $P = .16$	N=21 $P = .12$	N=20 $P = .009^*$	N=17 $P = .042^*$	N=13 $P = .22$	N=12 $P = .19$
Experiment 2	N=22 $P = .63$	N=22 $P = .78$	N=21 $P = .77$	N=21 $P = .46$	N=20 $P = .40$	N=18 $P = .52$	N=16 $P = .82$
Experiment 3	N=40 $P = .27$	N=40 $P = .19$	N=40 $P = .017^*$	N=40 $P = .006^*$	N=40 $P = .042^*$	N=38 $P = .036^*$	N=36 $P = .08$
Experiment 4	N=41 $P = .12$	N=40 $P = .24$	N=40 $P = .32$	N=40 $P = .36$	N=39 $P = .31$	N=38 $P = .15$	N=36 $P = .31$

References

1. Ball, W. (1973). The perception of causality in the infant. Presented at the Meeting of the Society for Research in Child Development, Philadelphia, PA.
2. Kotovsky, L., & Baillargeon, R. (2000). Reasoning about collisions involving inert objects in 7.5-month-old infants. *Developmental science*, 3(3), 344-359.
3. Wang, S. H., Kaufman, L., & Baillargeon, R. (2003). Should all stationary objects move when hit? Developments in infants' causal and statistical expectations about collision events. *Infant Behavior and Development*, 26(4), 529-567.
4. Muentener, P., & Carey, S. (2010). Infants' causal representations of state change events. *Cognitive psychology*, 61(2), 63-86.