For 19-Month-Olds, What Happens On the Screen Stays On the Screen

Barbu Revencu[1] **(revencu_barbu@phd.ceu.edu)** and **Gergely Csibra**[1, 2] **(csibrag@ceu.edu)**
[1] Department of Cognitive Science, Central European University
1051 Budapest, Hungary
[2] Department of Psychological Sciences, Birkbeck, University of London
WC1E 7HX London, UK

Abstract

Fictional entities in animations and puppet shows are widely used in infancy research, and there is plenty of evidence suggesting that infants are able to make inferences about them (e.g., ascribing agency to self-propelled 2-D figures). In the present set of experiments, we asked whether 19-month-olds take what they see on the screen to be happening in the here and now, or whether they think that on-screen events are spatiotemporally decoupled from the immediate environment. We found that infants do not expect an animated ball falling on a screen to end up in real boxes below the screen, even though they can track the ball (i) when the ball is real, and (ii) when the boxes are also part of the animation. These findings indicate that infants separate animations from the surrounding environment and cast doubt on the assumption that infants are naïve realists about iconic representations.

Keywords: representations; animation; development; fiction; methodology

Introduction

Humans stand out among other animals in their ability to go beyond their current environment and gather information about distal states of affairs from proximal sources (e.g., an utterance spoken *at present*, a screen *in front of me*). While an apple, a photo of an apple, and the Apple logo share *apple*-ness as part of their internal description, only the first of these affords *apple*-action (eating, cutting, peeling, etc.). That we navigate such stimuli so swiftly indicates an ability to decouple incoming percepts, which necessarily reach our senses here and now, from the information carried by those percepts (Ittelson, 1996; Millikan, 2017).

Take, for instance, Heider & Simmel's (1944) short animations of geometrical shapes moving around. When adults are asked to describe such clips, they respond as if they talked about real agents, attributing to them goals, desires, and intentions: the big triangle is chasing the small triangle, the circle wants to exit the enclosing, and the three shapes together form a love triangle (Heider & Simmel, 1944; Oatley & Yuill, 1985). Regardless, adults are not fooled into believing that these shapes *do* form romantic bonds: they know these are not fully-fledged agents, they are not afraid that the big bully triangle will chase them, and they do not consider interacting with the shapes. At least if prompted to think about it, adults would say that there is a designer behind the clip who used shapes and movement patterns to stand in for fictional agents and chasing events, respectively. In other words, they are aware that they are seeing a *representation of a fictional world*.

Similar stimuli are routinely used in developmental research to tap into early conceptual understanding and, in many cases, there is substantive evidence that young infants interpret them in an adult-like manner: they attribute instrumental and social goals to 2-D shapes (Gergely, Nádasdy, Csibra, & Bíró, 1995; Kuhlmeier, Wynn, & Bloom, 2003), they infer social relations from minimal interactions between these shapes (Powell & Spelke, 2013; Tatone, Geraci, & Csibra, 2015), and they ascribe mental states to animated caterpillars and flatfish (Surian, Caldi, & Sperber, 2007; Tauzin & Gergely, 2018). Undoubtedly, infants' inferences are prompted by the low-level cues that they would use to detect agents outside the lab, such as face-like features, self-propelled movement and contingent responsivity (see Opfer & Gelman, 2011, for a review). Little is known, however, on what infants make of these stimuli once the interpretive process has started. As a first step, we asked whether infants think that animated events are continuous with the surrounding reality.

We tested this question by investigating whether 19-month-olds expect a ball falling on the screen to land in boxes below the screen. First, we measured a baseline for infants' accuracy in tracking real balls falling in one of two boxes (Experiment 1, *Reality Baseline*). Second, we tested whether infants expect animated balls falling on the screen to land in boxes below the screen (Experiment 2, *Crossover*). Third, we ran a control experiment, in which both the ball and the boxes were part of the animation, to make sure that infants can follow an animated ball's trajectory when everything happens on the screen (Experiment 3, *Animation*). The three setups are illustrated in Figure 1. The hypotheses and dependent measures for all three experiments were preregistered at the Open Science Framework (here and here).

Figure 1: Overview of Three Experimental Setups: *Reality*, *Crossover*, *Animation* (left to right).

Experiment 1: Reality Baseline

Methods

Participants The final sample consisted of 16 typically developing 19-month-olds (M_{age} = 19 months 14 days, SD_{age} = 12.38 days).

Materials We built a wooden seesaw (height = 40 cm; width = 60 cm) that could be inclined left and right (angle $\approx$ 25°) by means of a 25-cm handle extending from the back of the seesaw, which allowed us to manipulate the seesaw from behind a curtain (Figure 1, left). We used several identical-looking red sponge balls (radius = 2.5 cm) and two different-colored rectangular cardboard boxes ($14 \times 15 \times 26$ cm^3) as landing positions for the balls dropped from the seesaw. Within each box, we added a dividing wall to create two compartments. This ensured that balls in the box were not accessible to infants even if they tried to open the boxes. The back compartments were padded with soft cloth to remove acoustic cues. In addition, we used two plush toys (a cat and a bird), which were hidden in the boxes to familiarize infants with the task of pointing to object locations, and a canvas bag for storing the toys and balls throughout the procedure.

Stimuli A small loudspeaker, placed behind the seesaw, played a 1-second jingle before each test trial, to prompt infants to attend to the ball-falling event.

Familiarization Infants entered the lab room with their caregivers and were seated on their caregivers' lap, approximately 40 centimeters from the experimental table (Figure 1, left). The experimenter drew the infant's attention to the two boxes, showed them that they can be opened, and revealed their (empty) insides. She then took a plush toy cat from a canvas bag and allowed the infant to inspect the toy for 10 seconds. Meanwhile, she pushed the inner compartments backwards, so she would be able to drop the toy into the boxes. She then asked the infant to hand the toy, moved behind the seesaw, drew the infant's attention to herself ("[Name,] look!"), and dropped the toy into one of the two boxes (*e.g.*, the orange box). She then slid the inner compartments back into place, pushed the boxes to the edge of the table, where the infant could reach them, and asked "Where is it?". If the infant failed to respond within 3 seconds, she asked them "Where is the cat?" two more times (at 10-second intervals) before retrieving the toy from the box herself. If infants touched or pointed to the right box, the experimenter congratulated the infant and took the toy out from the box. If infants picked the wrong one, the experimenter showed them that the box they chose is empty and retrieved the toy from the box where it had been dropped. The next familiarization trial was identical except that the cat was replaced by a toy bird. When infants responded correctly for 2 trials in a row, the experimenter put the toys away, pushed the boxes to the left and right of the seesaw, and pulled their inner compartments so the ball could fall into the boxes.

Test While looking at the ball from behind the seesaw, the experimenter drew the infant's attention to the red ball in the middle of the seesaw ("[Name,] look at the ball!"). Immediately afterwards, infants heard a 1-second jingle coming from a loudspeaker behind the seesaw and saw the ball falling either left or right into one of the two boxes. The experimenter did not follow the ball with her gaze, but kept her eyes on the middle of the seesaw. After the ball fell, the seesaw was brought back into horizontal position. The experimenter then pushed the boxes to the edge of the table, and asked the infant "Where is it?". Just like in familiarization, infants received two more prompts before ending the test trial. Unlike in familiarization, infants were given neutral feedback, by being congratulated regardless of their choice, and the ball was *not* removed from the box. Once the trial ended (when infants chose a box or 10 seconds after the third question), infants were handed one of the two toys from familiarization and encouraged to play with it while the experimenter set up the next trial (pushing the boxes back next to the seesaw and placing a new ball in the middle of the seesaw). Each infant received 4 test trials.

Design The location alternated across familiarization and test such that the toy in the last familiarization trial and the ball in the first trial always ended up in opposite boxes (AB-ABBA). The side with which the AB-ABBA alternation started (left vs. right), the side of the boxes (orange right, blue left vs. orange left, blue right), and the experimenter's position during the test question (to the left vs. to the right of the seesaw) were counterbalanced. For each trial, we measured whether the infant made a choice and, if they did, whether the box they chose was on the same side as the falling event.

Coding We had two primary dependent measures: choice and correctness. Infants received a score of 1 for having made a choice if they unambiguously reached, grasped, or pointed to one of the two boxes, and 0 otherwise. Infants' correctness was coded as 1 if they chose the box that was on the same side as the falling event, and as 0 otherwise. Infants' responses were recorded by one researcher during the testing session, and double-coded from video by a second researcher who was blind to the ball location. Inter-rater reliability was very high (Cohen's κ = 0.858); inconsistencies were solved by discussion.

Exclusion Criteria We excluded infants who did not make two correct choices in a row across 8 familiarization trials (n = 4). One additional infant was excluded due to experimenter error. In addition, we excluded trials in which infants did not follow the ball trajectory with their gaze based on video recordings (2 out of $16 \times 4 = 64$ trials). One additional trial was excluded due to experimenter error.

Results

Before proceeding to data analysis, infants' raw scores for each trial (0 or 1 for choice, 0 or 1 for correctness if infants made a choice) were converted into aggregate individual scores: the proportion of choices across trials, and the proportion of correct responses across the trials where a choice has been made. All analyses were conducted in R.

As expected, infants were able and motivated to solve the task. Most of them provided at least one response (87.5%, 14 out of 16 participants), and they did so in 72.1% of the trials (44 out of 61). When they made a choice, their responses were correct 83.3% of the time (median = 1, Wilcoxon signed rank[1], $V = 82.5$, $p = .007$, $r = .655$), well above the 50% chance level.

Discussion

The purpose of Experiment 1 was twofold: (i) to make sure that infants can follow the trajectory of balls falling into boxes; and (ii) to get a quantitative baseline of this capacity when the entire setup consists of real objects. The results indicate that 19-month-olds can answer questions about displaced objects reliably (70% choices) and accurately (83% correct choices). This benchmark allowed us to go for the main question of the study and investigate whether infants would do the same in a situation in which screen events appear to extend into the surrounding environment.

Experiment 2: Crossover

Methods

Participants The final sample consisted of 16 typically developing Hungarian 19-month-olds (M_{age} = 19 months 7 days, SD_{age} = 13.9 days).

Materials We used an LCD TV screen (16:9, diagonal 110 cm) to play animations, in which a ball on the screen fell either to the left or the right. The same boxes used in Experiment 1 were placed under the screen to create the illusion that the ball lands into them (Figure 1, center). In familiarization, we used the same two plush toys (a cat and a bird) as in Experiment 1.

Stimuli We transposed the events from Experiment 1 in a 2D-animated format, using *Adobe Animate CC*: a red ball (more precisely, a red circle) falling off a seesaw to the left or to the right (Figure 1, center). The dimensions of the animated ball and seesaw matched those of the real objects. The same jingle used in Experiment 1 was played from the TV speakers before the ball-falling event to draw infants' attention to the screen.

Familiarization The warmup phase was identical to Experiment 1: the experimenter dropped a toy into one of the two boxes and asked the infant where the toy was.

Test Test trials followed the same logic as those in Experiment 1. While behind the screen, the experimenter drew the infant's attention to the red ball on the screen ("[Name,] look at the ball!"), which then rolled to the left or to the right of the seesaw. The experimenter pushed the boxes away from the screen and towards the infant, and asked them "Where is it?". The trial ended if the infant chose one of the two boxes or if they did not respond to the third prompt. The experimenter handed one of the toys from familiarization, which they could play with while she set up the next trial. Each infant received 4 test trials.

Design The design was identical to Experiment 1 (AB-ABBA location alternation), and we counterbalanced the same factors (first location of the ball, side of the two boxes, and experimenter's position at test).

Coding We had the same two primary dependent measures, choice and correctness, as in Experiment 1. The responses were recorded by one researcher during the testing session, and double-coded from video by a second researcher who was blind to the side on which the ball had fallen. Inter-rater reliability was substantial (Cohen's $\kappa = .761$); inconsistencies were solved by discussion. Based on pilot data, we preregistered a secondary measure and coded how often infants pointed to the center of the screen when not choosing one of the two boxes.

Exclusion Criteria Just as in Experiment 1, we excluded infants who did not make two correct choices in a row across 8 familiarization trials ($n = 4$) and trials in which infants did not look at the falling event ($n = 2$). Two additional trials were excluded due to experimenter error.

Predictions If infants decouple the events on the screen from the surrounding environment, they should not expect the animated ball to cross the boundaries of the screen. We thus predicted that infants would (i) make fewer choices than in the *Reality Baseline*; and (ii) be at chance when making a choice.

Results

Unlike in Experiment 1, only 50% of the infants chose a box at least once during test (8 out of 16 participants). Out of the 60 valid trials included in the final analysis, infants picked out a box in 18 trials only (30%). Our secondary measure allowed us to rule out that infants were less motivated to provide an answer to the question in this version of the task:

[1] As 10 out of 16 infants were at ceiling (accuracy score of 1), accuracy scores were not normally distributed (Shapiro-Wilk, $W = .784$, $p = .002$), hence the use of non-parametric tests.

in 24 out of the remaining 42 trials (57%), infants pointed to the screen when asked where the ball was.

When they did make a choice, infants chose the box that was on the same side of the falling event 45.8% of the time (median = 0.5, Wilcoxon signed rank, $V = 3.5$, $p = .71$, $r = .196$).

Discussion

In the *Crossover* version of the falling ball experiment, infants behaved in a way that is inconsistent with the belief that animations are spatiotemporally continuous with reality. In contrast to their behavior in Experiment 1, they were less likely to choose a box when asked where the ball was, and often preferred to point to the screen. When they did provide a response, however, they did not base their answer based on the side of the falling event and chose a box at random instead.

Nonetheless, it is possible that infants simply did not get the intended referent of the question "Where is the ball?" because they did not see the red animated circle as a potential candidate for "ball", and that they pointed to the screen to request another animation. To rule out this alternative explanation, we added the two boxes to the animated world: if infants understand the question as we intended them to, they should now be able to point (again) to the correct location when asked about the ball's whereabouts.

Experiment 3: Animation

Methods

Participants The final sample consisted of 16 typically developing Hungarian 19-month-olds (M_{age} = 19 months 3 days, SD_{age} = 12.8 days).

Procedure As we sought to exclude the possibility that events happening on the screen are more difficult to process by infants, the materials, stimuli, procedure, and design were identical to Experiment 2 except for the boxes, which were now also part of the animation (Figure 1, right[2]).

Coding The two primary dependent measures, choice and correctness, were the same as in Experiment 1. Subjects' responses were recorded by one researcher during the testing session, and double-coded from video by a second researcher who was blind to the ball location. Inter-rater reliability was very high (Cohen's κ = 0.804); inconsistencies were solved by discussion.

Predictions We expected infants to be able to track ball displacement when everything happens within an animation. We thus predicted the same pattern of results as in the *Reality Baseline*: when infants choose, they would choose correctly (i.e., the box that is on the same side as the falling event).

Results

Comparable to the Reality Baseline, 81.3% of infants gave at least one response (13 out of 16 participants). Out of the 62 valid trials included in the final analysis, infants chose a box in 30 trials (48.4%). As for accuracy, infants chose the box that was on the same side of the falling event far from the 50% chance level: they pointed to the correct box in 93.6% of the trials in which they made a choice (median = 1, Wilcoxon signed rank, $V = 78$, $p < .001$, $r = .864$).

Discussion

While they made fewer choices overall compared to Experiment 1, infants overwhelmingly pointed to the box into which they saw the animated ball last fall on trials where they made a choice. This suggests that the random pattern of pointing in the *Crossover* Experiment was neither due to infants' inability to link the animated red circle to the intended referent of "the ball", which this experiment was designed to control for, nor due to other differences between Experiments 1 and 2 (e.g., the fact that the experimenter could not herself see the ball when it fell because she was standing behind the TV screen).

Overall Results

Correctness The experiment-wise differences between infants' accuracy rates (choosing the box into which they last saw the red ball fall) are straightforward. As Figure 2 shows, accuracy rates across Experiments 1-3 do not come from the same distribution (Kruskal-Wallis, $\chi^2(2) = 13.658$, $p = .001$).

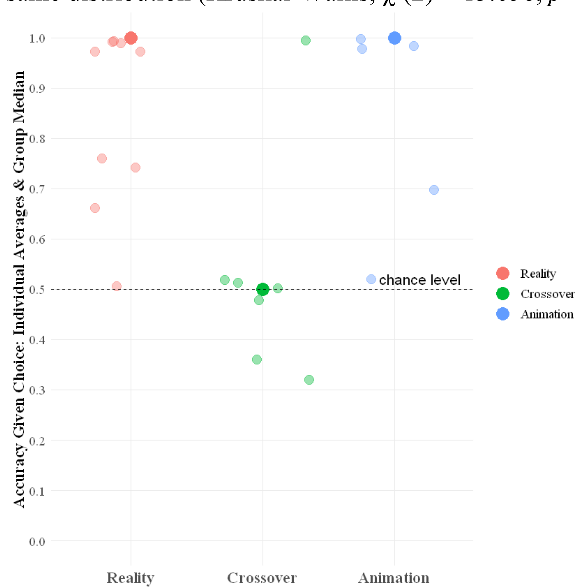

Figure 2: Accuracy rates (transparent: individual averages, opaque: group median) for trials in which infants chose a box.

[2] Once infants passed the familiarization phase with the two plush toys, the cardboard boxes were removed from the table.

The difference is driven by the *Crossover* Experiment, where infants were at chance between the two boxes (*Reality-Crossover*, Dunn's Test, z = 2.876, p = .008; *Animation-Crossover*, Dunn's Test, z = 3.612, p < .001; *Reality-Animation*, Dunn's Test, z = –.905, p = .366). When infants chose a box in Experiments 1 and 3, they chose it based on the falling event they had just seen. By contrast, in Experiment 2, they completely disregarded the animated falling event and randomly picked one of the two boxes.

Correctness and Choice To model both choice and accuracy rates, we built a Bayesian multinomial processing tree in JAGS (Plummer, 2003; Kruschke, 2014), which considers the two independent measures at once (Figure 3). Using infants' responses (*no choice*, *correct choice*, or *incorrect choice*), the model allows us to infer both (i) whether infants believe falling balls end up in boxes, and (ii) whether their beliefs differed across experiments.

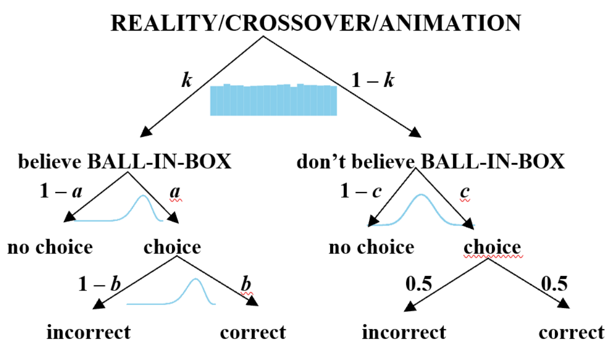

Figure 3: Hierarchical Bayesian model for all three experiments, modeling both choice and accuracy.

We use k (ranging from 0 to 1) to denote infants' beliefs about ball location in each experiment, and we make no a priori assumptions about what the values of k would be before seeing the data. We do, however, make two assumptions as to how beliefs and responses are linked. First, we assume that infants are more likely to make a choice *and* to choose correctly if they believe that the ball is in one of the two boxes (indicated by the mildly skewed priors on the left side of the tree). Second, we assume that infants are equally likely to refrain or to choose a box (at random) when they don't think that the ball is in either of the two boxes (as shown by the balanced priors on the right side of the tree).

Having constructed the data-generating model (from infants' beliefs to their responses), we use Bayes' rule to invert it to infer infants' beliefs from their responses to beliefs. In the extreme case, if infants always choose *and* choose correctly, they probably believe that the ball is in the box (left side of the tree). On the other hand, if infants make a choice only half of the time, and are at chance when choosing, they probably don't think that the ball is in the box (right side of the tree). Thus, large k-values (closer to 1) would indicate that infants believe there is a ball in the box into which they last saw it fall; conversely, small k-values (closer to 0) would indicate that infants do not entertain this belief.

The posteriors on the overarching parameter k (different for each experiment) as well as the posterior on the distribution of pairwise differences between Experiments are in line with frequentist analyses (Figure 4).

The first row in Figure 4 shows individual posterior distributions for all three experiments. For Experiment 1, k_{Reality} peaks around 1, suggesting that infants rely on the previous ball falling event when answering the test question. Similarly, $k_{\text{Animation}}$ also peaks towards the right end of the [0, 1] interval, but the estimate is noisier because infants made fewer choices than in Experiment 1. By contrast, $k_{\text{Crossover}}$ shows the opposite trend towards 0, indicating that infants do not think that the animated ball ends up in real boxes.

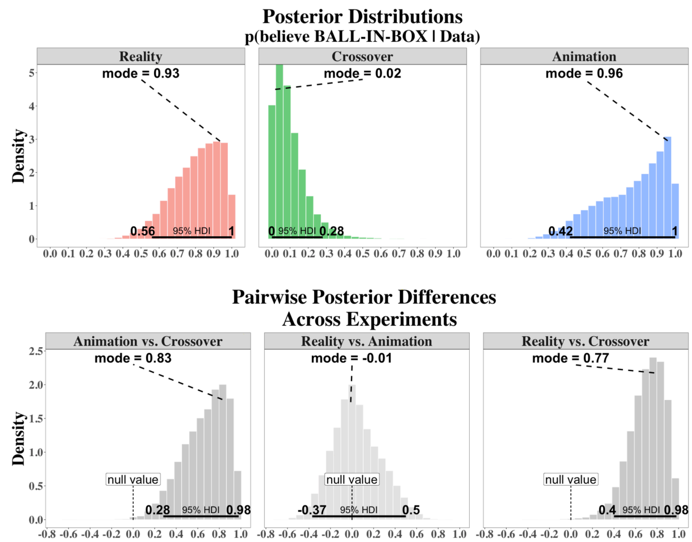

Figure 4: Posterior distributions for the k-parameter in each of the three experiments (top row), and posterior distributions on pairwise differences between Experiments (bottom row). Bold horizontal lines above the x-axis give the 95% highest density interval of the distributions.

The second row in Figure 4 computes the posterior of distribution differences for the three experiments. While the *Reality* and *Animation* distributions for the k-parameter are not very different from one another (mode of difference posterior is close to 0), the *Crossover* distribution diverges from the other two (the 95% highest density intervals for both difference posteriors exclude 0 as a plausible value).

General Discussion

Fiction and representations are relevant for developmental psychology because they are pervasively used to elicit infants' and children's inferences (e.g., animations, puppet shows, games). Strange setups involving fictional worlds transgressing boundaries are often used in developmental research under the assumption that infants are naïve realists. In a study by Lucca, Pospisil, & Sommerville (2018), for instance, 13-month-olds saw two people dropping objects on-screen, and were encouraged to choose one of the trays on the floor (as a measure of their preference between the two on-screen agents). Did infants really believe that these are the same objects they saw on the screen? At the heart of these methodologies lies the tacit assumption that infants take these stimuli at face value, which seems unlikely in the light of our experiments, at least for 19-month-olds. If our infant participants had been naïve realists, we would have seen the same choice patterns across all three experiments.

Taken together, the data indicate that 19-month-olds do not think that events on the screen extend beyond the screen. When the setup is fully real or fully animated, infants have no difficulty tracking the ball, from the seesaw into one of the two boxes, and indicate the right box when asked where the ball is (Experiment 1 & 3). By contrast, when infants watch an animation that appears to continue beyond the screen, infants are not fooled into thinking that the boundary can in fact be crossed (Experiment 2): infants either ignore the boxes and point to the screen, or pick one of them at random.

Conclusion

The world of infants is not a spatiotemporal hodgepodge. By 19 months of age, they have figured out that what happens on a screen stays on the screen, and they can answer questions about objects' locations appropriately based on this knowledge. However, it remains an open question whether infants view these animations as representations (of fictional or real events) or whether they have just learned that screens are *spatially* disconnected from their surroundings, while still believing that the events depicted on the screen are happening *now* (like in an aquarium). We are currently running a study meant to tease apart between these two possibilities.

We think it is important to figure out what infants make of fiction and representations for both theoretical and methodological reasons. First, the literature has focused predominantly on how children understand representations of particular states of affairs (see DeLoache, 2004, for a review; but see Preissler & Carey, 2007). Humans, however, gather plenty of information from representations that are not about immediate reality, so it is unclear why the former should come first in development. Second, this line of research is directly relevant to the methodology of developmental studies, where it is often tacitly assumed that fictional worlds and the real one will be treated as one and the same by infants and children.

Acknowledgments

We thank Iulia Savos and Dori Mészégető for data collection and coding, Ádám Koblinger for Bayesian guidance, and Laura Schlingloff and Denis Tatone for discussions.

This research has received partial funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement no. 742231 (PARTNERS).

References

DeLoache, J. S. (2004). Becoming symbol-minded. *Trends in cognitive sciences*, *8*(2), 66-70.

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*(2), 165-193.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. The *American Journal of Psychology*, *57*(2), 243-259.

Ittelson, W. H. (1996). Visual perception of markings. *Psychonomic Bulletin & Review*, *3*(2), 171-187.

Kruschke, J. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. Academic Press.

Kuhlmeier, V., Wynn, K., & Bloom, P. (2003). Attribution of dispositional states by 12-month-olds. *Psychological science*, *14*(5), 402-408.

Lucca, K., Pospisil, J., & Sommerville, J. A. (2018). Fairness informs social decision making in infancy. *PloS one, 13*(2), e0192848.

Millikan, R. G. (2017). *Beyond concepts: Unicepts, language, and natural information.* Oxford University Press.

Oatley, K., & Yuill, N. (1985). Perception of personal and interpersonal action in a cartoon film. *British journal of social psychology*, *24*(2), 115-124.

Opfer, J. E., & Gelman, S. A. (2011). Development of the animate-inanimate distinction. In U. Goswami (ed.), *The Wiley-Blackwell Handbook of Childhood Cognitive Development*, *Second Edition* (pp. 213-238). West Sussex, UK: Blackwell Publishers Ltd.

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd international workshop on distributed statistical computing* (Vol. 124, No. 125.10).

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.

Preissler, M., & Carey, S. (2004). Do both pictures and words function as symbols for 18-and 24-month-old children?. *Journal of Cognition and Development*, *5*(2), 185-212.

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological science*, *18*(7), 580-586.

Tatone, D., Geraci, A., & Csibra, G. (2015). Giving and taking: Representational building blocks of active resource-transfer events in human infants. *Cognition*, *137*, 47-62.

Tauzin, T., Gergely, G. (2018). Tauzin, T., & Gergely, G. (2018). Communicative mind-reading in preverbal infants. *Scientific reports*, *8*(1), 9534.