Observing others' joint attention increases 9-month-old infants' object encoding

Maleen Thiele[1,2], Robert Hepach[3], Christine Michel* [2,4] & Daniel B.M. Haun*[1,2]

[1]Max Planck Institute for Evolutionary Anthropology, Department of Comparative Cultural Psychology, Leipzig

[2]Leipzig Research Center for Early Child Development, Department of Early Child Development and Culture, Leipzig University

[3]Department of Experimental Psychology, University of Oxford

[4]Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig

*joint senior author

Manuscript is published in Developmental Psychology © 2021, American Psychological Association. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. Please refer to the published paper at https://doi.org/10.1037/dev0001189

Corresponding author:

Maleen Thiele

Department of Comparative Cultural Psychology

Max Planck Institute for Evolutionary Anthropology

Deutscher Platz 6, 04103 Leipzig, Germany

Phone: 0341-3550-407

Email: maleen_thiele@eva.mpg.de

Abstract

In direct interactions with others, 9-month-old infants' learning about objects is facilitated when the interaction partner addresses them through eye contact before looking toward an object. In this study we investigated whether similar factors promote infants' observational learning from third-party interactions. In Experiment 1, 9-month-old typically developing infants from mixed socioeconomic backgrounds from urban Germany (*N* = 32, 13 female) were presented with four types of videos showing one object and two adults. The scenarios varied systematically regarding the eye contact between the adults (eye contact or no eye contact), and the adults' object-directed gaze (looking toward or away from the object). To assess infants' encoding performance we measured their looking times when seeing the familiarized object subsequently next to a novel object, interpreting an enhanced novelty preference as reversely indicating greater encoding of the familiarized object. Infants showed an increased novelty preference, but only after observing a joint attentional setting during which two adults attended to the familiarized object together (following eye contact). In Experiment 2, we found an identical pattern of results in a matched first-party design in which 9-month-old infants (*N* = 32, 16 female) were directly addressed by one single adult on screen. Infants' encoding was only enhanced when the adult made eye contact with the infant before looking at an object. Together, this suggests that the capacity to learn through observing others' interactions emerges already in the first postnatal year, and that it may depend on similar factors as their learning through direct social engagement.

Keywords: object encoding, social attention, social interaction, infant development, ostension, eye tracking

Social interactions represent an essential source of learning opportunities in infancy. Both active participation in interactions as well as the observation of others' interactions help infants to acquire knowledge about their environment (Paradise & Rogoff, 2009; Rogoff, Paradise, Arauz, Correa-Chávez, & Angelillo, 2003). During the second half of the first postnatal year, infants' social behavior undergoes a crucial development as they begin to engage in triadic interactions. The transition from purely interpersonal dyadic (infant-adult) interactions to triadic (infant-adult-object) interactions represents a milestone in the first year of life, because it enables infants to incorporate external objects in their interactions and thereby communicate and learn about their environment (Tomasello, Carpenter, Call, Behne, & Moll, 2005). In triadic interactions with others, infants' learning about novel objects is strongly promoted when they see their interaction partner gazing in their direction before moving their attention toward an object (object-directed gaze). Both factors are vital components of *joint attention* interactions during which infant and adult coordinate their attention to an object of mutual interest (Carpenter, Nagell, Tomasello, Butterworth, & Moore, 1998). Underlying factors of infants' observational learning from *others'* interactions remain less clear. In the current study we investigated whether infants' learning through observation of others' interactions depends on similar factors as their learning in direct interactions.

Infants begin to engage in coordinated joint attention with others between 6 and 9 months of age (Carpenter et al., 1998; Striano & Bertin, 2005). In contrast to triadic *parallel* attention episodes in which infant and adult look to an object individually, triadic *joint* attention episodes include mutual awareness of the shared attentional experience. Early joint attention behavior is typically defined as the alternation of eye-gaze between two interaction partners and an object of mutual interest (e.g., Bakeman & Adamson, 1984). Breaking it down into its separate components, the joint attentional triangle includes three types of communicative looks (Carpenter & Liebal, 2011): An initiation look during which the initiator looks in the direction of the recipient to make eye contact, a reference look during which both partners gaze toward the object (in the following referred to as "object-directed gaze"), and a sharing look during which both partners re-engage in eye-contact, indicating their mutual awareness of the joint attentional experience.

During the same period as infants develop a joint attentional awareness, they begin to make *functional use* of joint attention in terms of learning about new objects. Studies with real-interactive settings have revealed that 7- and 9-month-old infants show increased object encoding when they are familiarized with a novel object in a joint attention situation as compared to a non-interactive control condition (Cleveland, Schug, & Striano, 2007; Cleveland & Striano, 2007). In the joint attention scenario, an experimenter established initial eye contact with the infant before alternating their gaze between a visible object and the infant. In the non-interactive "object-only" condition, the experimenter alternated their gaze between the object and a spot at the ceiling without looking in the direction of the infant at any point. As a measure of encoding, previous studies have used infants' response to novelty (Fantz, 1964; Rose, Melloy-Carminar, & Bridget, 1982). When seeing the familiarized object next to a novel object in a subsequent paired-preference test, 7- and 9-month-old infants (but not 4- and 5-month-olds) looked relatively longer to the novel object in the joint attention condition as compared to the object-only condition (Cleveland et al., 2007; Cleveland & Striano, 2007; Striano, Chen, Cleveland, & Bradshaw, 2006). Following the logic of the paradigm, this novelty preference reversely indicates that the familiarized object had been previously sufficiently encoded, causing higher attention to the novel, yet unprocessed object (for a review on infant visual attention and object recognition see Reynolds, 2015). In line with these findings, studies with slightly older children have shown that joint attention represents an important role (not a fundamental requirement, Scofield & Behrend, 2011; Tomasello, Strosberg, & Akhtar, 1996) for successful word learning around 2 years of age

(Baldwin, 1995; Hirotani, Stets, Striano, & Friederici, 2009; Tomasello & Farrar, 1986; see also Scofield & Behrend, 2011).

In focusing more specifically on others' initial eye contact and object-directed gaze as two components of the joint attentional triangle, screen-based eye tracking studies have revealed that the *interplay* between both factors has an important impact on infants' object encoding. Already at the early age of 4 months, infants show increased encoding of an object if they have seen an adult shifting their gaze toward this object. This is indicated by an increased response to a previously non-cued object (increased looking times in behavioral studies: Hoehl, Wahl, & Pauen, 2014; Reid & Striano, 2005; Wahl, Michel, Pauen, & Hoehl, 2013; enhanced neural response in physiological studies: Hoehl, Reid, Mooney, & Striano, 2008; Hoehl, Wahl, Michel, & Striano, 2012; Reid, Striano, Kaufman, & Johnson, 2004; Wahl et al., 2013). In addition to others' direct gaze to an object, infants' object encoding is influenced by ostensive signals in direct face-to-face interactions (Csibra & Gergely, 2006, 2009). At 9 months, infants' object encoding is only enhanced if the infant has been addressed before the object-directed gaze shift, for example, through direct gaze or infant directed speech. If the adult does not address the infant sufficiently before looking at the object, 9-month-olds can follow the adult's gaze to a target object (Gredebäck, Astor, & Fawcett, 2018; Szufnarowska, Rohlfing, Fawcett, & Gredebäck, 2015), but they do not show signs of encoding it (Okumura, Kanakogi, Kobayashi, & Itakura, 2020). The other way around, if the adult keeps looking in the direction of the child without shifting their gaze to a visible object, 9-month-old infants do not process the object, even if they fixate on it when the object location is highlighted non-socially by illumination (Okumura, Kanakogi, Kobayashi, & Itakura, 2017). Together this suggests that, in direct triadic interactions with others, 9-month-old infants' learning about an object is enhanced by the *interplay* between others' object-directed gaze and previous eye contact. This is in line with the Natural Pedagogy account, which suggests that ostensive signals like eye contact (but also infant-directed speech, calling of the infants' name, or contingent responses) increase infants' attention and responsiveness to learn-worthy content, and scaffold their referential learning in direct face-to-face interactions (Csibra & Gergely, 2006, 2009).

At least by the end of the second year of life, children can learn about objects by merely observing *others'* triadic interactions. Two-year-old toddlers, for example, learn novel object labels equally well when they are directly addressed as when they overhear a conversation between others (Akhtar, 2005; Akhtar, Jipson, & Callanan, 2001; Floor & Akhtar, 2006; O'Doherty et al., 2011). There is some indication from previous studies suggesting that this observational learning may be facilitated by similar processes as infants' learning from direct interactions. One recent study has, for example, shown that 2-year-old toddlers learned a novel object label only when observing two actors attending to an object together (i.e., following previous eye contact), but not when one of the actors was engaged with another activity (Fitch, Lieberman, Luyster, & Arunachalam, 2020). Moreover, 18-month-olds have been found to learn the location of a hidden object only if an adult had used ostensive cues to indicate the location for another adult (alternating their gaze between interaction partner and object location while pointing), not when the adult had used non-ostensive similar behaviors (absent-minded gazing and extended index finger, Gräfenhain, Behne, Carpenter, & Tomasello, 2009). These findings suggest that, at least by the age of 2 years, the sensitivity to ostensive cues may extend to third-party settings and that observed joint attention may promote observational learning about objects (Böckler, Knoblich, & Sebanz, 2011; see also Meng, Uto, & Hashiya, 2017).

Infants' sensitivity to third-party interactions develops well before their second birthday. Around 6 months of age, infants perform more saccadic gaze shifts in accordance with the flow of a conversation when two people are facing each other as compared to two people standing back-to-back while talking (Augusti, Melinder, & Gredebäck, 2010), they show increased pupil dilation in response to others' irrational compared to rational feeding interactions (Gredebäck & Melinder, 2010), and prefer to touch animated geometric-shape-agents who previously helped another agent over hindering characters (Hamlin, Wynn, & Bloom, 2007). Nine-month-old infants preferentially attend to face-to-face interactions over scenes in which the same two people turn away from one another (Handl, Mahlberg, Norling, & Gredebäck, 2013; see also Galazka, Roché, Nyström, & Falck-Ytter, 2014), and 10-month-olds expect a talking person to look in the direction of a human interaction partner (rather than an animated toy truck, Beier & Spelke, 2012; see also Bakker, Kochukhova, & von Hofsten, 2011). At 12 months of age, infants can predict the action goals in others' feeding interactions (Gredebäck & Melinder, 2010), and they anticipate a person to respond to another person's action only if this action serves a communicative goal (e.g., speech not coughing, Yamashiro & Vouloumanos, 2018; see also Thorgrimsson, Fawcett, & Liszkowski, 2015, 2014; von Hofsten, Uhlig, Adell, & Kochukhova, 2009). Together, these findings suggest that already in the second half of the first year of life infants can recognize and selectively attend to situations in which two people engage in a social interaction, keenly observe and visually examine these scenes, and begin to understand the communicative function of others' gaze and gestures without being addressed themselves. What remains unclear, however, is whether infants at this age can already make use of others' interactions to gather knowledge about their environment and, if so, whether similar processes as in direct interactions facilitate this observational learning. To directly compare the underlying processes of infants' learning from others' interactions with their learning from direct interactions, it requires a systematic investigation of the relative impact of third-party eye contact and third-party object-directed gaze on early observational learning.

The current study

The aim of this study was twofold: first, to investigate whether 9-month-old infants can learn about objects through mere observation of others' interactions and, second, to examine whether this learning depends on similar factors as infants' referential learning in direct interactions (i.e., the interplay between eye contact and object-directed gaze as present during joint attention). We tested 9-month-old infants because infants at this age typically engage systematically in joint attention behaviors themselves (e.g., Carpenter et al., 1998), and can make functional use of these situations in terms of object learning (e.g., Cleveland & Striano, 2007). We conducted two experiments allowing us to directly compare infants' object encoding in a third-party observational context (Experiment 1) with their encoding in a first-party context (Experiment 2). In Experiment 1, infants were presented with an object-encoding task during which they saw four types of videos each showing one object together with two adults (encoding phase). Based on a 2×2 design, we systematically varied the content of the videos with respect to the eye contact between the two actors (eye contact or no eye contact), as well as the actors' gazing at the object (looking at the object or looking away from the object). After each video, infants were presented with a preferential-looking phase during which they saw the object they had just seen (familiarized object) next to a novel object. As a measure of infants' previous object encoding performance, we calculated their proportional looking times to the novel object, interpreting an increased novelty preference as an indicator of increased processing of the familiar object (Reid & Striano, 2005). Based on previous studies, we hypothesized that if similar processes contribute to infants' learning in observational contexts as in direct interactions, infants' object encoding should be facilitated by observed joint attention, that is, when two actors gaze at an object together after previous eye contact. In this case, infants' proportional looking time to the novel object should be significantly higher in the third-party "eye contact – looking at object" condition compared to all other three conditions. If infants process information in observational contexts differently than in first-party contexts, the study design would allow to generate alternative explanations related to the relative impact of observed eye contact and others' gaze cues to an object. A main effect of one or both factors would, for example, indicate that infants' object encoding is influenced independently by observed eye contact and/or others' direct gazing at the object.

To compare infants' object encoding during third-party observation with a situation in which they are directly addressed, we ran a second experiment with a methodologically matched first-party design. In contrast to Experiment 1, the videos in Experiment 2 showed one single adult gazing in the direction of the infant (eye contact) or looking away from the infant (no eye contact) before looking toward (or away from) an object. Based on previous research, we expected that the interplay between eye contact and subsequent object-directed gaze would promote infants' object encoding (e.g., Cleveland & Striano, 2007; Okumura et al., 2020). More specifically, we hypothesized that infants' proportional looking time to the novel object should be significantly higher in the first-party "eye contact – looking at object" condition compared to the other three conditions.

Experiment 1

Methods

Ethical approval for the design and procedure of the study "Observing others' joint attention increases 9-month-old infants' object encoding" was provided by the Child Subjects Committee of the Max Planck Institute for Evolutionary Anthropology (no protocol number). We pre-registered the hypotheses, methods, procedures, and the data analysis plan for this experiment on the Open Science Framework (https://osf.io/yfegm/). Video examples, eye tracking raw data, and R scripts for pre-processing and analyzing the data are available at the same link on the Open Science Framework.

Participants. Thirty-two infants between 9 months, 0 days and 10 months, 0 days of age were included in the final sample of Experiment 1 (n = 13 female; M = 291.0 days, SD = 10.13 days). Data from three additional infants were excluded because they did not provide the minimum amount of one valid trial per condition (see section "Coding and Data Analysis" for valid trial criteria). All participants were born full-term (> 37 weeks). They were recruited on a voluntary basis via phone from the database of the Max Planck Institute for Evolutionary

Anthropology in Leipzig. Children in this data base come from Leipzig (Germany) or surrounding areas, an urban Central-European, industrialized context. We did not collect individual data regarding the participants' socioeconomic background, but families in this database typically come from mixed, mainly mid to high socioeconomic backgrounds. Written informed consent was obtained from one parent of each infant prior to testing. The participants received a small gift as thank you for their participation in the study. The sample size was planned based on a priori power analysis simulation including data from a pilot study (for details see section S3 in the supplemental materials).

Procedure. The testing took place at the Max Planck Institute for Evolutionary Anthropology in Leipzig (Germany). All participants were presented with an eye-tracking task, during which they sat in front of a screen on their parent's lap. We used a 23.8'' monitor with 93 dpi and 1920×1080 screen resolution. To run the experiment and to record infants' gaze movements, a Tobii eye tracker (Tobii X120, Tobii Technology, Stockholm, Sweden) and Tobii eye tracking software (Tobii Studio version 3.4.8.1348) was used. Data was recorded separately for the left and the right eye at a sampling frequency of 120 Hz. We used a five-point calibration procedure to calibrate the eye-tracker to the participant's eyes. The total duration of the Experiment was approximately 10 minutes, the whole visit at the lab around 20 minutes.

Stimuli and Design. All participants were presented with 16 trials of an object-encoding task. Each trial consisted of an object encoding phase and a preferential-looking phase (see also Wahl et al., 2013). In the beginning of each phase, a blinking star was presented in the center of the screen together with attention-grabbing sound to attract the infant's attention (2 seconds duration before the object encoding phase, and 1second duration before the preferential-looking phase). The overall duration of each trial was 24 seconds (see Figure 1).

Object encoding phase. Infants saw a video showing an object positioned between two adults. As objects, we used 32 pictures of abstract toys from the object stimuli collection applied in the study by Wahl and colleagues (2013). Based on a 2×2 design, the content of the videos was manipulated with regard to (a) the presence of initial eye contact between the two actors (eye contact or no eye contact), and (b) whether or not the actors looked at the object (looking toward or away from the object). To manipulate the eye contact between the actors, we used the movements of their bodies (turning toward or away from one another), the relative positioning of their bodies (face-to-face or back-to-back), and gaze direction (eye contact or looking in opposite directions). The actors kept a neutral facial expression and remained silent during the entire sequence (see also Meng et al., 2017). To clearly highlight the third-party context, the actors never faced in the direction of the participant. All videos were presented without sound. The four conditions resulting from the 2×2 design were tested within-subjects, with each child being presented with four trials of each condition (see also Szufnarowska et al., 2015).

The videos had a total duration of 11 seconds, with identical timing over all conditions and videos (see Figure 1): Initially, the actors were seen in back view (1s) before they turned toward (or away from) one another (1s), and remained in this face-to-face or back-to-back position (1s). Then, both actors turned their heads and gaze simultaneously in the direction of (or away from) the object (1s) and remained in this position (5s). Finally, the actors turned their heads and gazes back (1s) and remained in the initial face-to-face or back-to-back position (1s). The videos were consistent with previous studies regarding the total duration, the non-social baseline sequence in the beginning, the duration of the actor's fixation times on the object, and the overall duration of eye contact between the interaction partners (Meng et al., 2017; Okumura, Kanakogi, Kanda, Ishiguro, & Itakura, 2013; Okumura et al., 2017, 2020; Szufnarowska et al., 2015; Theuring, Gredebäck, & Hauf, 2007). In contrast to the video stimuli used in previous studies, we split the two-second overall duration of the eye contact sequence in two one-second lasting phases. This way, the joint attention scenario ("eye contact – looking at object" condition) met the minimum requirement for a complete joint attentional triangle (Carpenter & Liebal, 2011), including initial eye contact, subsequent looking toward an object, as well as a "closing" eye contact sequence following the mutual look at the object (similar to the dynamic in previous studies with real-interactive settings, e.g., Cleveland et al., 2007). All videos were presented in full-screen view, covering an area of 48° width × 27° height (at a screen distance of 60 cm).

Every child saw two different dyads of actors: one dyad performing in all trials of the two social interactive conditions, and the other dyad performing in all trials of the two no eye contact conditions. The left-right positioning of the actors within the dyads was reversed in the "looking at object" and "no looking at object" conditions (see Figure 1). The two dyads were seen equally often in all conditions. The actors were all female, wore white t-shirts and were visible from the waist up. All possible body and head movements from all actors in all conditions covered an area of 14.5°×20.8° on both sides of the object, with the head movements covering an area of 11.2°×8.8°. The objects covered an area of 6.5°×6.5°. The minimum distance between each actor's face and the object was 9.3°. In section S1 in the supplemental materials we provide detailed information on how the video stimuli were created.

Preferential-looking phase. We presented infants with two objects at the same time: the object they had previously seen in the encoding phase (familiar object) and a novel object. The objects were presented side-by-side on a grey background for a duration of 10 seconds. The size of the objects was identical to the object size in the encoding phase. The distance between the outer edges of the objects was 18.2°. The positioning of the novel object (right or left) was counterbalanced within infants and condition (i.e., each infant saw the novel object in two trials per condition on the right side, and in two trials on the left side). The pairing of objects in the preferential-looking phase was randomized and consistent over participants. All infants saw the same 32 objects, with the same toy never occurring twice. Sixteen objects were presented as familiar objects in the object encoding phase, and 16 additional objects as novel objects in the preferential-looking phase. The role of each individual object (novel or familiar) was counterbalanced across participants and conditions, meaning that each object served equally often as novel and as familiar object in all four conditions over all participants.

The 16 experimental trials were presented in four blocks with four trials each. Within a given block, each trial presented a different condition (see also Szufnarowska et al., 2015). We counterbalanced the order of conditions during the first block across infants (in the way that an equal number of infants saw condition 1 first, condition 2 first, etc.). The order of conditions in the remaining three blocks was pseudo-randomized. No condition occurred more than twice in a row. After every block, infants were presented with a 4-second kaleidoscope video with a soothing melody to maintain their attention (see also Reuter, Emberson, Romberg, & Lew-Williams, 2018; Szufnarowska et al., 2015).

Coding and Data Analysis. We used the R software environment (R version 3.6.3, RStudio version 1.2.1335) for pre-processing and analyzing the data, as well as for setting areas of interest (AOIs). As main dependent variable, we measured infants' mean proportional looking time to the novel object in the preferential-looking phase. For this purpose, we defined two square-shaped AOIs: one AOI covering the novel object, and one AOI covering the familiar object. To accommodate for inaccuracies in calibration, all AOIs were defined 1° visual angle larger than the maximal dimensions of the stimuli (Gredebäck, Johnson, & von Hofsten, 2009). We assessed the total duration of fixations in both AOIs for each trial and participant, including fixation data from the entire preferential-looking phase. Data for both the left and the right eye of each participant was averaged. When one eye could not be measured, we used the data from the other eye. To define fixations, we used the Tobii Velocity-Threshold Identification (I-VT) fixation filter with default settings (for details see section S2 in the supplemental materials online). In a next step, we calculated the proportional looking time at the novel object ("novelty preference score") by dividing the duration of fixations to the novel object by the total duration of fixations to both objects. The novelty preference score could take values between 0 and 1, with values above .5 indicating a relatively longer looking time at the novel object. We only included a trial if infants had looked at least at one object during the preferential-looking phase for at least one fixation, and if they had paid visual attention to the central parts of the video for at least one fixation (see also Michel, Pauen, & Hoehl, 2017). In accordance with our pre-registered plan, we counted the initial face-to-face (or back-to-back) phase, and the looking-to-object (or away-from-object) phase as 'central parts', excluding the motion sequences. Infants were only included in the analysis if they provided valid data in at least one trial per condition after being filtered according to these criteria (see also Wahl et al., 2013). On average, infants provided 3.46 valid trials (*SD* = .69) per condition (for detailed valid trial statistics see Table S1 in the supplemental materials).

To test our hypotheses, we fitted a generalized linear mixed model (GLMM) with a Gaussian error structure. All models were fitted using the R package "lme4" (Version 1.1-21, Bates, Maechler, Bolker, & Walker, 2020). The dependent measure was the novelty preference score. To investigate infant's object encoding as a function of third-party eye contact and third-party object-directed gaze, we fitted a GLMM including the interaction between third-party eye contact (eye contact, no eye contact) and others' object-directed gaze (looking toward the object, not looking at object) as a fixed effect. To account for possible trial effects, trial (z-transformed) was included as control variable in the model. All factors were tested within-subjects. As random effects, we included subject as intercept, as well as random slopes on subject for trial, eye contact, object-directed gaze, and the interaction between eye contact and object-directed gaze. The significance of the individual fixed effects was based on likelihood ratio tests comparing the full models with respective reduced models excluding the individual fixed effects using the *drop1*-function in R with an alpha-level of .05.

We ran the following analyses in addition to the preregistered plan. First, we conducted six pair-wise comparisons to compare the novelty preference score between all four conditions. All pairwise comparisons were based on the GLMM fitted for the main analysis, using the R package "emmeans" (Version 1.4.6, Lenth, 2020). To account for multiple comparisons, the alpha-level was adjusted via Bonferroni correction. Second, we calculated one-sample t-tests against .50 within each condition to determine whether the novelty preference score significantly differed from chance level. In addition, we explored whether infants' direct attention to the object in the encoding phase could explain their looking preference in the subsequent preferential-looking phase. For this purpose, we fitted two additional GLMMs to investigate infants' novelty preference score (dependent variable). In one model, we included infants' looking time to the object in the encoding phase as fixed effect (within-subject factor). In the second model, we included a binary fixed effect variable indicating for each trial whether the child had looked at the object during encoding at all, that is, whether their fixation duration within the object AOI was greater than zero (within-subject factor: yes, no). Table S1 in the supplemental materials provides an overview of the corresponding valid trial statistics. In both models, we included subject as random intercept, as well as the random slope of the fixed effect on subject. Fixation data from the entire video sequence was included for the measure of looking time. We have conducted some further analyses to better understand the impact of infants' overt attention during the encoding phase on their encoding performance. We did not find any systematic condition differences in infants' looking times or gaze patterns. For conciseness, we present the corresponding results in section S2 in the supplemental materials online.

Results

The comparison between the full model and the reduced model revealed a significant result indicating that at least one of the fixed effects had an impact on infants' mean proportional looking time to the novel object in the preferential-looking phase ($\chi^2(4) = 14.99$, $p = .005$). More specifically, the interaction between third-party eye contact and third-party looking at the familiarized object had a significant effect on infants' novelty preference ($\chi^2(1) = 4.03$, $p = .04$, estimate = −.08, *SE* = .04; see Figure 3a). We did not find an effect of trial ($\chi^2(1) = 1.22$, $p = .27$, estimate = .01, *SE* = .01).

Additional analyses revealed that infants' novelty preference score in the joint attention condition ("eye contact − looking at object") was significantly higher compared to the scores in all other three conditions (see Table 1). Moreover, the score was significantly higher than chance level in only this condition ($M = .62$, $SD = .13$, $t(31) = 5.28$, $p < .001$, $d = .93$, see Table 2). Infants' preferential orienting to the novel object did not depend on their attention to the familiarized object in the encoding phase: Neither did infants' looking time to the familiarized object predict their subsequent novelty preference ($\chi^2(1) = .49$, $p = .48$, estimate = .01, *SE* = .01), nor was the preference score systematically influenced by whether or not infants had looked at the object during encoding at all ($\chi^2(1) = .14$, $p = .71$, estimate = −.009, *SE* = .02). Table S3 in the supplemental materials shows the descriptive statistics of the looking times during the encoding phase in all four conditions.

Discussion

We found that 9-month-old infants showed an increased looking preference for a novel over a familiarized object after they had observed two adults attending to the familiarized object jointly (i.e., following mutual eye contact). This novelty preference was significantly higher compared to situations in which infants observed two adults looking at each other but away from the object (purely interpersonal attention), two adults looking at the object individually in a back-to-back setting (non-coordinated parallel attention), or scenarios in which two adults looked neither at the object nor at each other (neither person- nor object-directed attention). Based on the assumption that novelty preference reflects better stimulus encoding (Cleveland et al., 2007) our findings suggest that, in a purely observational setting, others' joint attention toward an object can increase 9-month-olds' encoding of this object. We did not find any evidence supporting the assumption that infants' increased encoding depended on infants' overt attention to the object during encoding. This is in line with previous studies showing that direct attention to an object is not required for its encoding (Cleveland & Striano, 2007; Wahl et al., 2013). In the General Discussion we provide a more detailed discussion of this assumption.

To directly compare infants' object encoding performance in a third-party observational setting to a situation in which they were directly addressed, we tested an additional sample of infants in a second experiment investigating infants' object encoding in a matched first-party setting during which they saw one single adult on screen.

Experiment 2

Methods

The experimental design, procedure, as well as data pre-processing and analysis procedures were identical to Experiment 1, but this time only one actor was visible in the videos in the encoding phase. We pre-registered the hypotheses, methods, procedures, and the data analysis plan for this experiment on the Open Science Framework (https://osf.io/dp5cg/). Video examples, eye tracking raw data, and R scripts for pre-processing and analyzing the data are available at the same link on the Open Science Framework. The design and procedure of this experiment was approved by the same Ethics Committee as in Experiment 1.

Participants. Thirty-two full-term infants between 9 months, 0 days and 10 months, 0 days of age were included in the final sample of Experiment 2 (n = 16 female; M = 282.69

days, *SD* = 8.4 days). Data from four additional infants were excluded because they did not provide the minimum amount one valid trial per condition. The criteria for data inclusion were the same as in Experiment 1. The participants were recruited from the same data base as described in Experiment 1. Each child participated only in one of the two Experiments.

Stimuli and Design. The videos in the encoding phase were as similar as possible to previous first-party studies (e.g., Okumura et al., 2013, 2020), while keeping methodological and visual details consistent with Experiment 1 (see section S1 in the supplemental materials for more details regarding the video stimuli). Analogous to the 2×2 design in Experiment 1, the content of the videos in the encoding phase was manipulated with regard to (a) the gaze direction of the adult in relation to the infant (eye contact, or no eye contact), and (b) whether or not the actor looked at the object (looking toward or away from the object). To manipulate the actor's infant-directed gaze, we used the movement of the actor's body (turning toward the child or to the side), the relative positioning of the body (facing the child or averted), and gaze direction (looking toward the child or to the side).

The videos in this experiment were edited in a way that they had exactly the same timing and degree of motion as the videos in Experiment 1, except that they contained one and not two actors (see Figure 2): Initially, the actor was seen in back view (1s) before turning in the direction of the child or to the side (1s), and remaining in this position (1s). Then, the actor turned head and gaze simultaneously in the direction of (or away from) the object (1s), and remained in this position (5s). Finally, the actor turned her head and gaze back (1s), and remained again in this initial position (1s). All possible body and head movements from all actors in all conditions covered an area of 15°×20.8° on both sides of the object, with the head movements covering an area of 12.3°×8.8°. The minimum distance between each actor's face and the object was 8.4° (see Figure S2 in the supplemental materials for an illustration of the visual arrangement on screen). As in Experiment 1, each child was presented with four trials of each condition, presented in four blocks. Every child saw each of the 4 conditions performed by a different actor. The actors were the same as in the first Experiment. They were shown equally often in all four conditions, counterbalanced between infants. The positioning of the actor left or right from the central object was counterbalanced within condition, child, and block (i.e., each child saw the actor in each block during two trials on the right side and during two trials on the left side; over the whole experiment, the actor occurred two times on the right side and two times on left side within each condition).

Coding and Data Analysis. Data pre-processing and trial inclusion criteria were the same as in Experiment 1. Importantly for the manipulation of eye contact in Experiment 2, this ensured that only trials were included in which the infant had payed attention to the initial direct-gazing phase during which the adult looked in the direction of the infant. We conducted the same main analyses and post-hoc tests. Analogous to Experiment 1, our main dependent variable was infants' proportional looking time to the novel object in the preferential looking phase. In the main model, we included the interaction between eye contact (eye contact, no eye contact) and object-directed gaze (looking toward or away from the object) as main fixed effect. The same control variables and random effects were included as in Experiment 1. On average, infants provided 3.42 valid trials (SD = .81) per condition (see Table S2 in the supplemental materials for the detailed valid trial statistics).

Results

The full-null model comparison revealed a significant result ($\chi^2(4) = 15.64$, $p = .004$). Specifically, the interaction between eye contact and others' object-directed gaze had a significant impact on infants' novelty preference score ($\chi^2(1) = 5.77$, $p = .02$, estimate = −.11, SE = .04; see Figure 3b). We did not find an effect of trial ($\chi^2(1) = 1.47$, $p = .22$, estimate = −.01, SE = .01).

Additional analyses revealed that infants' novelty preference in the "eye contact − looking at object" condition was significantly higher compared to all other three conditions (see Table 3). Moreover, the score was significantly higher than chance level in only this condition (M = .61, SD = .13, t(31) = 4.62, p < . 001, d = .82, see Table 4). Infants' preference for the novel object did not depend on the time they had looked at the object during the encoding phase ($\chi^2(1)$ = .003, p = .95, estimate = −.001, SE = .01), or on whether they had looked at the object during this phase at all ($\chi^2(1)$ = 1.03, p = .30, estimate = −.04, SE = .04). Table S4 in the supplemental materials shows the descriptive statistics of the looking times during the encoding phase in all four conditions.

Merged analysis over both Experiments. To explore possible differences between Experiment 1 and 2, we repeated our main analysis over a merged sample including infants from both experiments (N = 64). In addition to the fixed effects included in the main analysis, we included Experiment (1 or 2) as fixed effect. As in the separate analyses of the two experiments, we found that infants' novelty preference varied as a function of eye contact and looking at the familiar object (effect of the interaction: $\chi^2(1)$ = 8.35, p = .004, estimate = −0.09, SE = 0.03). We did not find any effect of trial ($\chi^2(1)$ = 0.0001, p = .99, estimate = 0.00, SE = 0.008) or experiment ($\chi^2(1)$ = 1.97, p = .16, estimate = −0.02, SE = 0.02). Identical to the separate analyses of each experiment, infants preferred the novel over the familiar object only in the "eye contact – looking at object" condition when combining data of both experiments.

Discussion

We found that 9-month-old infants showed an increased looking preference for a novel over a familiarized object if the adult had established eye contact with the infant before looking at the familiar object. This novelty preference was relatively higher compared to situations in which infants had seen an adult looking in their direction but away from the object, an adult looking at the object but not at the infant, or an adult looking neither at the object nor the infant. Infants' increased novelty preference could not be explained by increased overt attention to the object during encoding. Our findings suggest that the interplay between others' object-directed gaze and previous eye contact promotes object learning in 9-month-old infants, as the two factors alone were not sufficient to elicit this effect. This represents a conceptual replication of previous studies showing that joint visual attention affects infants' subsequent looking behavior and object encoding (e.g., Cleveland et al., 2007; Itakura, 2001; Okumura et al., 2020; Wahl et al., 2013). Additional analyses over a merged sample including participants from Experiment 1 and 2 suggested that infants' encoding performance in a third-party observational setting was identical to their performance in a first-party setting.

General Discussion

Previous studies have shown that, by the end of the second year of life, infants can learn about their environment by merely observing others' social interactions. In the present study we show that this ability emerges already in the second half of the first year after birth, and that early observational learning may be influenced by similar factors as infants' referential learning through direct interactions around this age. In Experiment 1, we found that 9-month-old infants showed an increased preference for a novel object compared to a familiarized object if they had previously seen two adults looking at the familiarized object in a joint attentional setting (i.e., following previous eye contact). This novelty preference was greater than chance level and higher compared to all other conditions, reversely indicating that observed joint attention to an object selectively increased infants' encoding of this object. In Experiment 2, we found the corresponding result pattern in a matched first-party design during which infants were directly addressed by one single adult on screen. Infants' novelty preference was highest if they had seen an adult addressing them through direct gaze before looking at the familiar object.

Our finding that the *interplay* between eye contact and others' object-directed gaze selectively increased infants' object encoding when being directly addressed, aligns with the theoretical assumption that communicative cues (Csibra & Gergely, 2006, 2009) and coordinated attention with others (Tomasello et al., 2005) provide an important framework for young infants' learning. Our finding that also *observed* joint attention can increase infants' object encoding, provides a crucial extension of this view by demonstrating that the promoting effect of joint attention goes beyond first-hand experience in direct pedagogical settings (see also Fitch et al., 2020 for supporting evidence with older children). Moreover and importantly, our findings imply that infants' successful learning through observation depended on *active* monitoring of others' interactions, not on passive reception (Oudeyer & Smith, 2016; Rogoff et al., 2003; see also Mani & Ackermann, 2018): When two adults established a communicative context by making eye contact, but then turned their gaze away from a visible object, infants did not show any signs of increased encoding of this object, suggesting that they invested their resources selectively in presumably meaningful triadic interactions. To enable selective learning, infants had to detect, selectively and intently observe, and evaluate the relevant learning opportunity. It remains to be studied what other factors besides observed coordination of visual attention define a meaningful social interaction. Following evidence from first-party studies, other potential factors could relate to characteristics of the interaction partners, such as parents versus strangers (e.g., Hoehl et al., 2012), neutral versus emotional affective states (e.g., Hoehl, Wiese, & Striano, 2008), or real versus artificial interaction partners (e.g., O'Doherty et al., 2011).

From a developmental perspective, the strikingly similar result pattern in the first- and third-party settings of our experiments is particularly noteworthy considering the narrow age range we tested in this study. The age around 9 months has been previously shown to represent a period during which infants' social development and learning undergo crucial developments as they develop competencies for triadic interactions (Carpenter et al., 1998; Striano & Rochat, 1999). By 9 months of age, infants can reliably coordinate their attention with others in joint attention episodes, and make functional use of these situations to acquire knowledge about external objects in their environment (Cleveland & Striano, 2007; Okumura et al., 2020). Our finding that, at the same age, the promoting effect of joint attention extends to merely observational settings, provides a novel perspective and important contribution to our understanding of the diverse pathways through which infants can learn from social interactions. Together with the previous literature on infants' active social interaction behavior, our findings raise diverse possibilities regarding the developmental relationship between the first- and third-party level. One possibility would be a causal relation. For example, infants' increasing own triadic skills and experiences during the second half of the first year of life may modulate their perception, understanding, and learning from others' joint attention or vice versa (see also Gredebäck & Melinder, 2010; Henderson, Wang, Matz, & Woodward, 2013; Herold & Akhtar, 2008). Alternatively, it would be possible that infants' sensitivity to own and others' joint attentional interactions develops in a parallel manner, driven by underlying social motivational processes (Dawson et al., 2004). To disentangle the relationship and possible causality of the two levels, longitudinal study designs would be needed.

Importantly, the main focus in this study was on the *outcome* of infants' encoding, measured by infants' visual recognition memory after the processing situation itself. To gain an additional insight into concrete processes during the actual encoding, we explored infants' looking pattern during the encoding phase further (see also section S2 in the supplemental materials for a discussion of our additional analyses). Overall, we did not find any indication that infants' object encoding would depend on their own direct attention to the object. This is in line with previous studies suggesting that overt attention to an object is not required for its processing (neurological studies: Hoehl, Wahl, et al., 2014; Reid & Striano, 2005; Reid et al., 2004; behavioral studies: Wahl et al., 2013; Okumura et al. 2013; Cleveland & Striano, 2007; Cleveland, 2007). Instead, it is likely that increased covert attentional orienting (i.e., shifts of visual attention occurring independently of eye movements) modulated infants' processing of the target object (Johnson, 1994; Posner, 1980). The exact processes of infants' covert attentional orienting during third-party joint attention remain unclear. One possibility would be that observed communicative cues such as eye contact highlight the meaningfulness of an upcoming referential gaze shift, causing a covert attentional shift in the direction of the referenced target object (Wahl et al., 2013; see also Daum & Gredebäck, 2011). Alternatively, the opportunity to observe a meaningful interaction may increase infants' responsiveness more generally, providing them with the necessary attentional activation to process everything within the range of the covert attentional field ("socially aware mode", Puce & Bertenthal, 2015). To disentangle these possibilities and to fully capture the mechanisms responsible for facilitated object encoding, it requires neural measures capturing attentional phenomena over and above direct visual attention. Moreover, infants' neural processes during the encoding situation need to be studied in relation to their subsequent behavioral and neuronal response to the object.

The implications of our findings are particularly profound with regard to cross-cultural differences in infants' everyday learning environments (Akhtar & Jaswal, 2020). Children across cultural contexts participate to highly variable extents in direct face-to-face interactions (e.g., Mesman et al., 2018). Infants in many cultures are thus much more accustomed to observational learning than to direct teaching (Paradise & Rogoff, 2009; Rogoff et al., 2007). From this perspective, our findings support the idea that joint attention per se, that is, independent of whether experienced through active participation or via observation of others, may represent a culturally universal communicative context in which generic knowledge can be transmitted (see also Correa-Chávez & Roberts, 2012). To investigate this possibility directly, it would require studies systematically comparing infants from different cultural backgrounds regarding their sensitivity and responsiveness to ostensive signals in direct interactions (e.g., Hernik & Broesch, 2019) as well as in third-party settings (e.g., Correa-Chávez & Rogoff, 2009).

Limitations

The findings of this study need to be considered against some limitations. First, our definition of joint attention (and, thus, the manipulation in our study design) was based on gaze cues. This was intended because mutual eye contact and others' object-related gaze have been previously identified as important influential factors on infants' object encoding. Moreover, gaze alternation between an object and an interaction partner has been used as an early indicator of 9-month-olds' joint attentional awareness in studies on infants' active social behavior (e.g., Bakeman & Adamson, 1984). To account for the broader range of social cues that infants encounter and produce in real interactions, future studies should test the generalizability of our findings to other ostensive signals (e.g., infant directed speech, calling the infants' name, contingent responsivity; Csibra, 2010), and other referential social cues such as pointing or showing gestures (especially once infants begin to use these gestures themselves to initiate joint attention, Carpenter et al., 1998; Mundy et al., 2007). This relates to another restriction of our findings, namely the highly controlled stimuli and testing environment. This limitation is particularly relevant regarding the interpretation of the similar result pattern in the first-party and third-party context. To investigate whether, in the real world, third-party joint attention is equally powerful on infants' learning as actively shared attention, our finding needs to be probed in a more natural environment. Moreover, looking-time based measures alone cannot directly inform us about the immediate underlying mechanisms that promote infants' learning in active interactive and observational settings (Aslin, 2007). Future studies are required using additional measures, such as measures of neural or physiological activity to capture signs of underlying mechanisms like (emotional) arousal (see, e.g., Hepach & Westermann, 2016 for a review on pupillometry in infancy research).

Importantly, our focus was on influential factors within the *social* part of the joint attentional triangle, not on the *kind* of information that infants learn about novel objects within different contexts. In focusing more specifically on the latter aspect, previous studies have suggested that 9-month-old infants may retain qualitatively different information about novel objects in direct communicative as compared to non-communicative contexts (Okumura, Kobayashi, & Itakura, 2016; Yoon, Johnson, & Csibra, 2008). While communicative context induced the representation of information about object identity (relevant for generalization) at the cost of learning information about object location, infants preferentially encoded the opposite information in non-communicative settings (Yoon et al., 2008; but see Silverstein, Gliga, Westermann, & Parise, 2019 for failed replication attempts). Investigating the influence of such a communication-induced memory bias in third-party settings would deepen our understanding of the kind of information infants learn from observing others' interactions, and how this compares to their learning in direct pedagogical settings.

Conclusion

In summary, we could show that observing others' interactions can increase 9-month-old infants' object encoding, and that this early observational learning is influenced by similar factors as infants' learning from direct social interactions at the same age. In both situations, the interplay between eye contact and object-directed gazing selectively enhanced infants' encoding of novel information. These findings suggest that not only active participation in joint attention, but also the mere observation of others' joint attention has a profound impact on early social learning. At a broader level, this has significant implications for the understanding of the multifaceted ways in which human infants learn about their world.

References

Akhtar, N. (2005). The robustness of learning through overhearing. *Developmental Science*, *8*, 199–209. https://doi.org/10.1111/j.1467-7687.2005.00406.x

Akhtar, N., & Jaswal, V. K. (2020). Stretching the Social: Broadening the Behavioral Indicators of Sociality. *Child Development Perspectives*, *14*, 28–33. https://doi.org/10.1111/cdep.12351

Akhtar, N., Jipson, J., & Callanan, M. A. (2001). Learning Words through Overhearing. *Child Development*, *72*, 416–430. https://doi.org/10.1111/1467-8624.00287

Aslin. (2007). What's in a look?. *Developmental science, 10*. 48-53. https://doi.org/10.1111/j.1467-7687.2007.00563.x

Augusti, E.-M., Melinder, A., & Gredebäck, G. (2010). Look who’s talking: Pre-verbal infants’ perception of face-to-face and back-to-back social interactions. *Frontiers in Psychology*, *1*. https://doi.org/10.3389/fpsyg.2010.00161

Bakeman, R., & Adamson, L. B. (1984). Coordinating Attention to People and Objects in Mother-Infant and Peer-Infant Interaction. *Child Development*, *55*, 1278. https://doi.org/10.2307/1129997

Bakker, M., Kochukhova, O., & von Hofsten, C. (2011). Development of social perception: A conversation study of 6-, 12- and 36-month-old children. *Infant Behavior and Development*, *34*, 363–370. https://doi.org/10.1016/j.infbeh.2011.03.001

Baldwin, D. A. (1995). Understanding the link between joint attention and language. Joint attention: Its origins and role in development, 131, 158.

Barton, K. (2019). *Mu-MIn: Multi-model inference.* R Package version 1.40.0. https://cran.r-project.org/web/packages/MuMIn

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2020). *Lme4: Linear mixed-effects models using “Eigen” and S4 classes.* R package version 1.1.15. https://cran.r-project.org/web/packages/lme4/lme4.pdf

Begus, K., Southgate, V., & Gliga, T. (2015). Neural mechanisms of infant learning: Differences in frontal theta activity during object exploration modulate subsequent object recognition. *Biology Letters*, *11*, 20150041. https://doi.org/10.1098/rsbl.2015.0041

Beier, J. S., & Spelke, E. S. (2012). Infants' Developing Understanding of Social Gaze: Infants' Developing Understanding of Social Gaze. Child Development, 486-496. https://doi.org/10.1111/j.1467-8624.2011.01702.x

Böckler, A., Knoblich, G., & Sebanz, N. (2011). Observing shared attention modulates gaze following. *Cognition*, *120*, 292–298. https://doi.org/10.1016/j.cognition.2011.05.002

Carpenter, M., & Liebal, K. (2011). Joint attention, communication, and knowing together in infancy. *Joint Attention: New Developments in Psychology, Philosophy of Mind, and Social Neuroscience*, 159–181.

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social Cognition, Joint Attention, and Communicative Competence from 9 to 15 Months of Age. *Monographs of the Society for Research in Child Development*, *63*, i. https://doi.org/10.2307/1166214

Cleveland, A., Schug, M., & Striano, T. (2007). Joint attention and object learning in 5- and 7-month-old infants. *Infant and Child Development*, *16*, 295–306. https://doi.org/10.1002/icd.508

Cleveland, A., & Striano, T. (2007). The effects of joint attention on object processing in 4- and 9-month-old infants. *Infant Behavior and Development*, *30*, 499–504. https://doi.org/10.1016/j.infbeh.2006.10.009

Correa-Chávez, M., & Roberts, A. L. D. (2012). A cultural analysis is necessary in understanding intersubjectivity. *Culture & Psychology*, *18*, 99–108. https://doi.org/10.1177/1354067X11427471

Correa-Chávez, M., & Rogoff, B. (2009). Children's attention to interactions directed to others: Guatemalan mayan and european american patterns. *Developmental Psychology*, *45*, 630–641. https://doi.org/10.1037/a0014144

Csibra, G. (2010). Recognizing Communicative Intentions in Infancy. *Mind & Language*, *25*, 141–168. https://doi.org/10.1111/j.1468-0017.2009.01384.x

Csibra, G., & Gergely, G. (2006). Social learning and social cognition: The case for pedagogy. *Processes of Change in Brain and Cognitive Development. Attention and Performance XXI*, *21*, 249–274.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*, 148–153. https://doi.org/10.1016/j.tics.2009.01.005

Daum, M. M., & Gredebäck, G. (2011). The development of grasping comprehension in infancy: Covert shifts of attention caused by referential actions. *Experimental Brain Research*, *208*, 297–307. https://doi.org/10.1007/s00221-010-2479-9

Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., & Liaw, J. (2004). Early Social Attention Impairments in Autism: Social Orienting, Joint Attention, and Attention to Distress. *Developmental Psychology*, *40*, 271–283. https://doi.org/10.1037/0012-1649.40.2.271

Fantz, R. L. (1964). Visual Experience in Infants: Decreased Attention to Familiar Patterns Relative to Novel Ones. *Science*, *146*, 668–670. https://doi.org/10.1126/science.146.3644.668

Fitch, A., Lieberman, A. M., Luyster, R. J., & Arunachalam, S. (2020). Toddlers' word learning through overhearing: Others' attention matters. Journal of experimental child psychology, 104793. https://doi.org/10.1016/j.jecp.2019.104793

Floor, P., & Akhtar, N. (2006). Can 18-Month-Old Infants Learn Words by Listening In on Conversations? *Infancy*, *9*, 327–339. https://doi.org/10.1207/s15327078in0903_4

Galazka, M. A., Roché, L., Nyström, P., & Falck-Ytter, T. (2014). Human Infants Detect Other People's Interactions Based on Complex Patterns of Kinematic Information. *PLoS ONE*, *9*, e112432. https://doi.org/10.1371/journal.pone.0112432

Gräfenhain, M., Behne, T., Carpenter, M., & Tomasello, M. (2009). Young children's understanding of joint commitments. *Developmental Psychology*, *45*, 1430–1443. https://doi.org/10.1037/a0016122

Gredebäck, G., Astor, K., & Fawcett, C. (2018). Gaze Following Is Not Dependent on Ostensive Cues: A Critical Test of Natural Pedagogy. *Child Development*, *89*, 2091–2098. https://doi.org/10.1111/cdev.13026

Gredebäck, G., Johnson, S., & von Hofsten, C. (2009). Eye Tracking in Infancy Research. *Developmental Neuropsychology*, *35*, 1–19. https://doi.org/10.1080/87565640903325758

Gredebäck, G., & Melinder, A. (2010). Infants' understanding of everyday social interactions: A dual process account. *Cognition*, *114*, 197–206. https://doi.org/10.1016/j.cognition.2009.09.004

Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution, 7*, 493–498. https://doi.org/10.1111/2041-210X.12504

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, *450*, 557–559. https://doi.org/10.1038/nature06288

Handl, A., Mahlberg, T., Norling, S., & Gredebäck, G. (2013). Facing still faces: What visual cues affect infants' observations of others? *Infant Behavior and Development*, *36*, 583–586. https://doi.org/10.1016/j.infbeh.2013.06.001

Henderson, A. M. E., Wang, Y., Matz, L. E., & Woodward, A. L. (2013). Active Experience Shapes 10-Month-Old Infants' Understanding of Collaborative Goals. *Infancy*, *18*, 10–39. https://doi.org/10.1111/j.1532-7078.2012.00126.x

Hepach, R., & Westermann, G. (2016). Pupillometry in infancy research. *Journal of Cognition and Development*, *17*, 359–377.

Hernik, M., & Broesch, T. (2019). Infant gaze following depends on communicative signals: An eye-tracking study of 5- to 7-month-olds in Vanuatu. *Developmental Science*, *22*, e12779. https://doi.org/10.1111/desc.12779

Herold, K. H., & Akhtar, N. (2008). Imitative learning from a third-party interaction: Relations with self-recognition and perspective taking. *Journal of Experimental Child Psychology*, *101*, 114–123. https://doi.org/10.1016/j.jecp.2008.05.004

Hirotani, M., Stets, M., Striano, T., & Friederici, A. D. (2009). Joint attention helps infants learn new words: Event-related potential evidence: *NeuroReport*, *20*, 600–605. https://doi.org/10.1097/WNR.0b013e32832a0a7c

Hoehl, S., Reid, V. M., Mooney, J., & Striano, T. (2008). What are you looking at? Infants' neural processing of an adult's object-directed eye gaze. *Developmental Science*, *11*, 10–16. https://doi.org/10.1111/j.1467-7687.2007.00643.x

Hoehl, S., Wahl, S., Michel, C., & Striano, T. (2012). Effects of eye gaze cues provided by the caregiver compared to a stranger on infants' object processing. *Developmental Cognitive Neuroscience*, *2*, 81–89. https://doi.org/10.1016/j.dcn.2011.07.015

Hoehl, S., Wahl, S., & Pauen, S. (2014). Disentangling the Effects of an Adult Model's Eye Gaze and Head Orientation on Young Infants' Processing of a Previously Attended Object. *Infancy*, *19*, 53–64. https://doi.org/10.1111/infa.12035

Hoehl, S., Wiese, L., & Striano, T. (2008). Young Infants' Neural Processing of Objects Is Affected by Eye Gaze Direction and Emotional Expression. *PLoS ONE*, *3*, e2389. https://doi.org/10.1371/journal.pone.0002389

Ishikawa, M., Yoshimura, M., Sato, H., & Itakura, S. (2019). Effects of attentional behaviours on infant visual preferences and object choice. *Cognitive Processing*, *20*, 317–324. https://doi.org/10.1007/s10339-019-00918-x

Itakura, S. (2001). Attention to repeated events in human infants (Homo sapiens): Effects of joint visual attention versus stimulus change. *Animal Cognition*, *4*, 281–284. https://doi.org/10.1007/s100710100087

Johnson, M. H. (1994). Visual attention and the control of eye movements in early infancy. *Attention and performance XV: Conscious and nonconscious information processing, 15,* 291-310.

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2020). *emmeans: estimated marginal means.* R package version 1.4. 4. https://cran.r-project.org/web/packages/emmeans/emmeans.pdf

Mani, N., & Ackermann, L. (2018). Why Do Children Learn the Words They Do? *Child Development Perspectives*, *12*, 253–257. https://doi.org/10.1111/cdep.12295

Meng, X., Uto, Y., & Hashiya, K. (2017). Observing Third-Party Attentional Relationships Affects Infants' Gaze Following: An Eye-Tracking Study. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.02065

Mesman, J., Minter, T., Angnged, A., Cissé, I. A. H., Salali, G. D., & Migliano, A. B. (2018). Universality Without Uniformity: A Culturally Inclusive Approach to Sensitive Responsiveness in Infant Caregiving. *Child Development*, *89*, 837–850. https://doi.org/10.1111/cdev.12795

Michel, C., Pauen, S., & Hoehl, S. (2017). Schematic eye-gaze cues influence infants' object encoding dependent on their contrast polarity. *Scientific Reports*, *7*. https://doi.org/10.1038/s41598-017-07445-9

Mundy, P., Block, J., Delgado, C., Pomares, Y., Van Hecke, A. V., & Parlade, M. V. (2007). Individual Differences and the Development of Joint Attention in Infancy. *Child Development*, *78*, 938–954. https://doi.org/10.1111/j.1467-8624.2007.01042.x

O'Doherty, K., Troseth, G. L., Shimpi, P. M., Goldenberg, E., Akhtar, N., & Saylor, M. M. (2011). Third-Party Social Interaction and Word Learning From Video: Social

Interaction and Video. *Child Development*, *82*, 902–915. https://doi.org/10.1111/j.1467-8624.2011.01579.x

Okumura, Y., Kanakogi, Y., Kanda, T., Ishiguro, H., & Itakura, S. (2013). The power of human gaze on infant learning. *Cognition*, *128*, 127–133. https://doi.org/10.1016/j.cognition.2013.03.011

Okumura, Y., Kanakogi, Y., Kobayashi, T., & Itakura, S. (2017). Individual differences in object-processing explain the relationship between early gaze-following and later language development. *Cognition*, *166*, 418–424. https://doi.org/10.1016/j.cognition.2017.06.005

Okumura, Y., Kanakogi, Y., Kobayashi, T., & Itakura, S. (2020). Ostension affects infant learning more than attention. *Cognition*, *195*, 104082. https://doi.org/10.1016/j.cognition.2019.104082

Okumura, Y., Kobayashi, T., & Itakura, S. (2016). Eye Contact Affects Object Representation in 9-Month-Old Infants. *PLOS ONE*, *11*, e0165145. https://doi.org/10.1371/journal.pone.0165145

Oudeyer, P.-Y., & Smith, L. B. (2016). How Evolution May Work Through Curiosity-Driven Developmental Process. *Topics in Cognitive Science*, *8*, 492–502. https://doi.org/10.1111/tops.12196

Paradise, R., & Rogoff, B. (2009). Side by Side: Learning by Observing and Pitching In. *Ethos*, *37*, 102–138. https://doi.org/10.1111/j.1548-1352.2009.01033.x

Posner, M. I. (1980). Orienting of attention. *Quarterly journal of experimental psychology, 32,* 3-25. https://doi.org/10.1080/00335558008248231

Puce, A., & Bertenthal, B. I. (Eds.). (2015). *The many faces of social attention: Behavioral and neural measures*. Cham , Heidelberg New York Dodrecht London: Springer.

Reid, V. M., & Striano, T. (2005). Adult gaze influences infant attention and object processing: Implications for cognitive neuroscience. *European Journal of Neuroscience*, *21*, 1763–1766. https://doi.org/10/bp4dcf

Reid, V. M., Striano, T., Kaufman, J., & Johnson, M. H. (2004). Eye gaze cueing facilitates neural processing of objects in 4-month-old infants. *Neuroreport*, *15*, 2553–2555.

Reuter, T., Emberson, L., Romberg, A., & Lew-Williams, C. (2018). Individual differences in nonverbal prediction and vocabulary size in infancy. *Cognition*, *176*, 215–219. https://doi.org/10.1016/j.cognition.2018.03.006

Reynolds, G. D. (2015). Infant visual attention and object recognition. *Behavioural Brain Research*, *285*, 34–43. https://doi.org/10.1016/j.bbr.2015.01.015

Rogoff, B., Moore, L., Najafi, B., Dexter, A., Correa-Chávez, M., & Solís, J. (2007). Children's Development of Cultural Repertoires through Participation in Everyday Routines and Practices. In *Handbook of socialization: Theory and research.* (pp. 490–515). New York, NY, US: The Guilford Press.

Rogoff, B., Paradise, R., Arauz, R. M., Correa-Chávez, M., & Angelillo, C. (2003). Firsthand Learning Through Intent Participation. *Annual Review of Psychology*, *54*, 175–203. https://doi.org/10.1146/annurev.psych.54.101601.145118

Rose, S. A., Gottfried, A. W., Melloy-Carminar, P., & Bridger, W. H. (1982). Familiarity and novelty preferences in infant recognition memory: Implications for information processing. *Developmental Psychology, 18,* 704. https://doi.org/10.1037/0012-1649.18.5.704

Scofield, J., & Behrend, D. A. (2011). Clarifying the role of joint attention in early word learning. *First Language*, *31*, 326–341. https://doi.org/10.1177/0142723710395423

Senju, A., & Csibra, G. (2008). Gaze Following in Human Infants Depends on Communicative Signals. *Current Biology*, *18*, 668–671. https://doi.org/10.1016/j.cub.2008.03.059

Silverstein, P., Gliga, T., Westermann, G., & Parise, E. (2019). Probing communication-induced memory biases in preverbal infants: Two replication attempts of Yoon, Johnson and Csibra (2008). *Infant Behavior and Development*, *55*, 77–87. https://doi.org/10.1016/j.infbeh.2019.03.005

Striano, T., & Bertin, E. (2005). Coordinated affect with mothers and strangers: A longitudinal analysis of joint engagement between 5 and 9 months of age. *Cognition & Emotion*, *19*, 781–790. https://doi.org/10.1080/02699930541000002

Striano, T., Chen, X., Cleveland, A., & Bradshaw, S. (2006). Joint attention social cues influence infant learning. *European Journal of Developmental Psychology*, *3*, 289–299. https://doi.org/10.1080/17405620600879779

Striano, T., & Rochat, P. (1999). Developmental link between dyadic and triadic social competence in infancy. *British Journal of Developmental Psychology*, *17*, 551–562. https://doi.org/10.1348/026151099165474

Szufnarowska, J., Rohlfing, K. J., Fawcett, C., & Gredebäck, G. (2015). Is ostension any more than attention? *Scientific Reports*, *4*, 5304. https://doi.org/10.1038/srep05304

Theuring, C., Gredebäck, G., & Hauf, P. (2007). Object processing during a joint gaze following task. *European Journal of Developmental Psychology*, *4*, 65–79. https://doi.org/10.1080/17405620601051246

Thiele, Michel, Hepach, & Haun (2020). *Infants' Object Processing in the Context of Third-Party Social Interactions.* The Open Science Framework. https://osf.io/yfegm/

Thiele, Michel, Hepach, & Haun (2020). *The Role of Ostensive Context and Gaze Cues on Infants' Object Processing.* The Open Science Framework. https://osf.io/dp5cg/

Thorgrimsson, G. B., Fawcett, C., & Liszkowski, U. (2014). Infants' expectations about gestures and actions in third-party interactions. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00321

Thorgrimsson, G. B., Fawcett, C., & Liszkowski, U. (2015). 1- and 2-year-olds' expectations about third-party communicative actions. *Infant Behavior and Development*, *39*, 53–66. https://doi.org/10.1016/j.infbeh.2015.02.002

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, *28*, 675–691.

Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child development,* 1454-1463. https://doi.org/10.2307/1130423

Tomasello, M., Strosberg, R., & Akhtar, N. (1996). Eighteen-month-old children learn words in non-ostensive contexts. *Journal of Child Language*, *23*, 157–176. https://doi.org/10.1017/S0305000900010138

von Hofsten, C., Uhlig, H., Adell, M., & Kochukhova, O. (2009). How children with autism look at events. *Research in Autism Spectrum Disorders*, *3*, 556–569. https://doi.org/10.1016/j.rasd.2008.12.003

Wahl, S., Michel, C., Pauen, S., & Hoehl, S. (2013). Head and eye movements affect object processing in 4-month-old infants more than an artificial orientation cue. *British Journal of Developmental Psychology*, *31*, 212–230. https://doi.org/10.1111/bjdp.12001

Yamashiro, A., & Vouloumanos, A. (2018). How do infants and adults process communicative events in real time? *Journal of Experimental Child Psychology*, *173*, 268–283. https://doi.org/10.1016/j.jecp.2018.04.011

Yoon, J. M. D., Johnson, M. H., & Csibra, G. (2008). Communication-induced memory biases in preverbal infants. *Proceedings of the National Academy of Sciences*, *105*, 13690–13695. https://doi.org/10.1073/pnas.0804388105

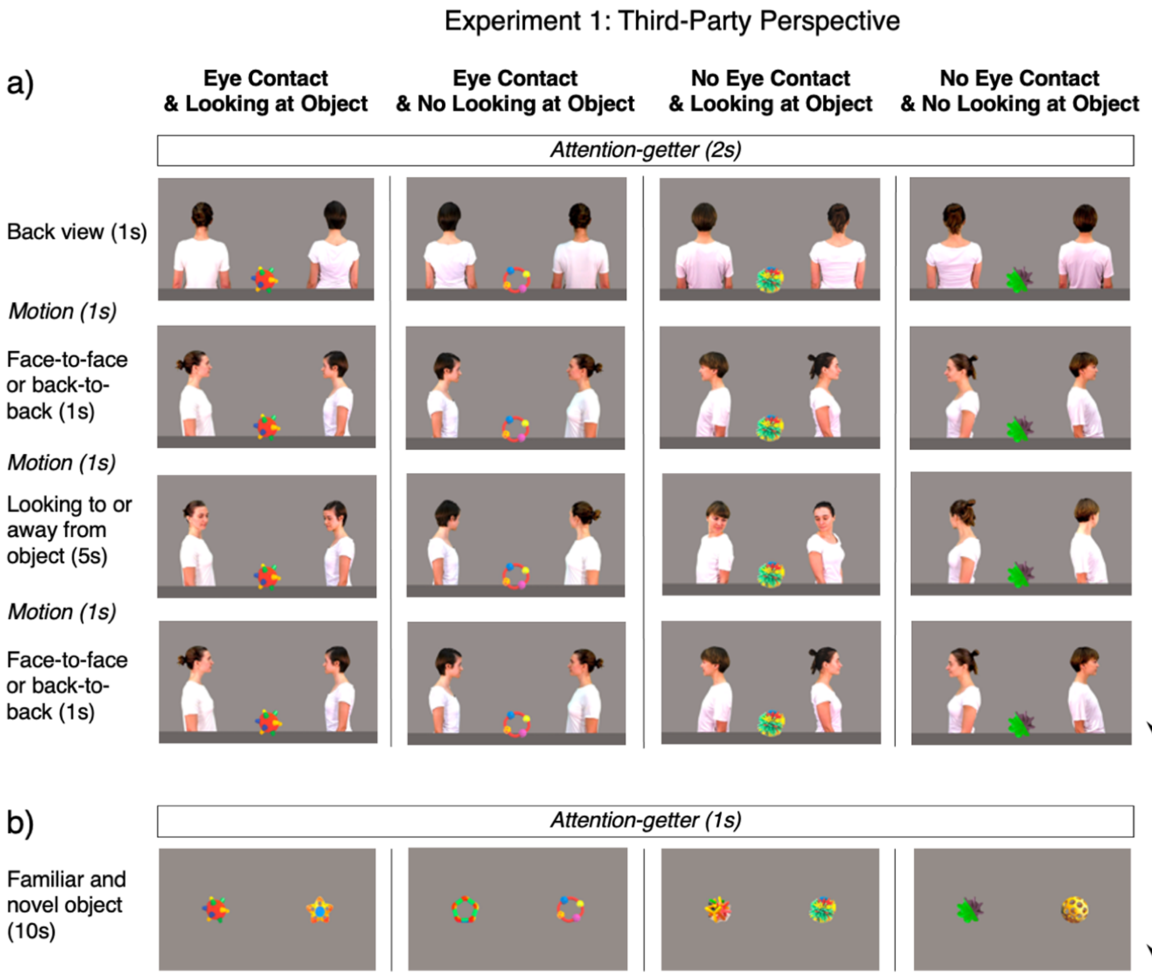

Figure 1. Exemplary sequence and timing of one trial of the object processing task for each of the four experimental conditions in Experiment 1. Every trial consisted of (a) an object encoding phase (11s), and (b) a preferential-looking phase (10s). Before each phase, an attention-getter (blinking star) was presented in the center of the screen on a black background. The actors depicted in this illustration gave signed consent to be published in this article.

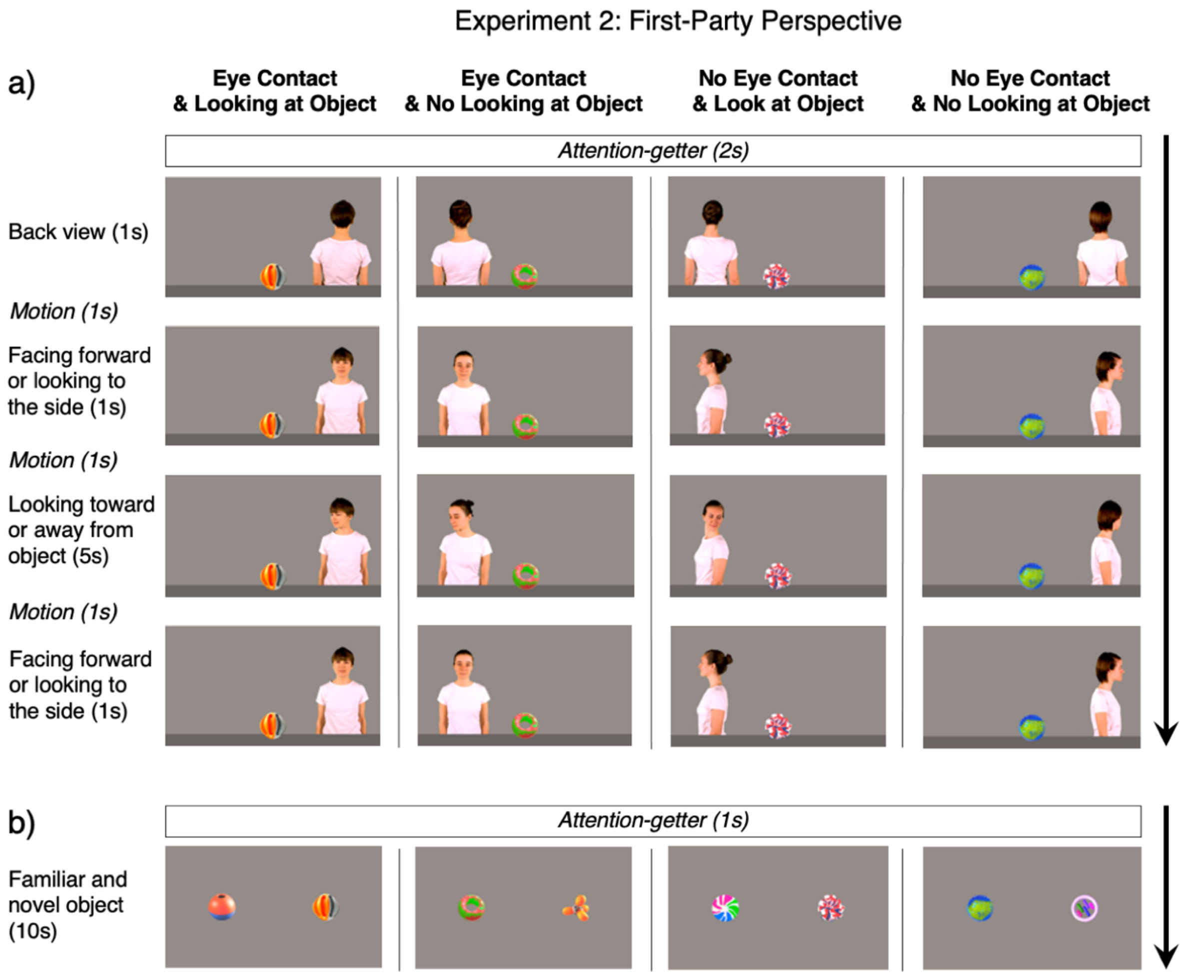

Figure 2. Exemplary sequence and timing of one trial of the object processing task for each of the four experimental conditions in Experiment 2. Every trial consisted of (a) an object encoding phase (11s) and (b) a preferential-looking phase (10s). Before each phase, an attention-getter (blinking star) was presented in the center of the screen on a black background. The actors depicted in this illustration gave signed consent to be published in this article.

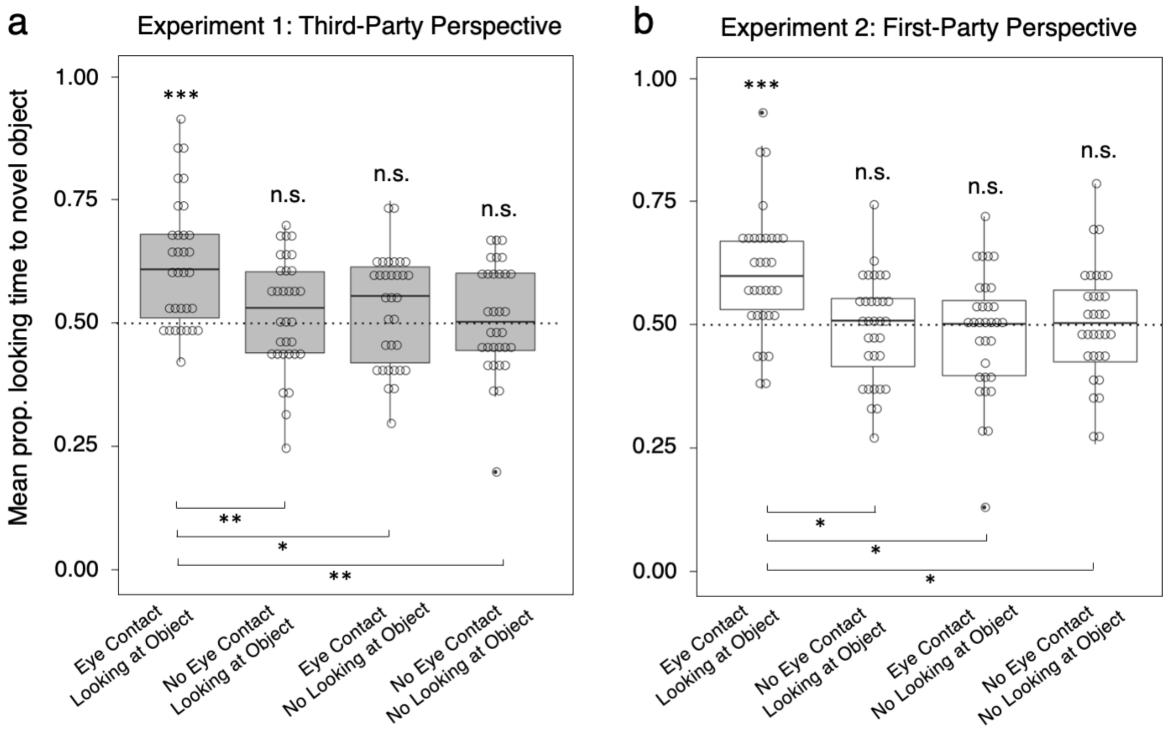

Figure 3. Boxplots with individual data points illustrating the mean proportional looking time to the novel object in the preferential-looking phase for the four conditions of (a) Experiment 1 and (b) Experiment 2. The dashed line at .50 represents chance level. The results of the post-hoc one-sample t-tests against chance level are depicted above the boxplots, the results of the pair-wise comparisons below. P-values of the pair-wise comparisons are adjusted via Bonferroni correction. *$p < .05$ **$p < .01$ ***$p < .001$

Table 1

Results from post-hoc pair-wise comparisons of the novelty preference score between the four conditions of Experiment 1 (Third-Party Perspective)

Compared conditions	estimate	*SE*	*df*	*t*	*p*
Eye Contact – Looking at Object vs. No Eye Contact – Looking at Object	–.10	.03	81.9	–3.28	.009**
Eye Contact – Looking at Object vs. Eye Contact – No Looking at Object	.08	.03	74.8	2.89	.03*
Eye Contact – Looking at Object vs. No Eye Contact – No Looking at Object	.10	.03	91.8	3.40	.006**
Eye Contact – Looking at Object vs. Eye Contact – No Looking at Object	–.01	.03	89.0	–0.48	1.0
No Eye contact – Looking at Object vs. No Eye Contact – No Looking at Object	.002	.03	160.2	0.08	1.0
Eye Contact – No Looking at Object vs. No Eye Contact – No Looking at Object	–.02	.03	150.9	–0.56	1.0

Notes. Results are based on the pairwise contrasts between the estimated marginal means of all conditions, inferred from the main GLMM fitted to the data by using the R-package "emmeans". P-values are adjusted via Bonferroni correction for six tests.

Table 2

Results from post-hoc one-sample t-tests testing the novelty preference score within the four conditions of Experiment 1 against chance

Condition	M*(SD)*	*t*	*df*	*p*	*d*
Third-Party Eye Contact Looking at object	.62*(.13)*	5.28	31	<.001***	.93
No Third-Party Eye Contact Looking at object	.52*(.11)*	0.96	31	.34	.17
Third-Party Eye Contact No Looking at object	.53*(.11)*	1.47	31	.15	.26
No Third-Party Eye Contact No Looking at Object	.51*(.11)*	0.49	31	.62	.09

Table 3

Results from post-hoc pair-wise comparisons of the novelty preference score between the four conditions of Experiment 2 (First-Party Perspective)

Compared conditions	estimate	*SE*	*df*	*t*	*p*
Eye Contact – Looking at object vs. No Eye Contact – Looking at object	–.10	.03	79.3	–3.17	.01*
Eye Contact – Looking at object vs. Eye Contact – No looking at object	.10	.03	79.3	3.23	.01*
Eye Contact – Looking at object vs. No Eye Contact – No looking at object	.10	.03	92.9	3.05	.02*
No Eye Contact – Looking at object vs. Eye Contact – No looking at object	.002	.03	91.4	0.07	1.0
No Eye Contact – Looking at object vs. No Eye Contact – No looking at object	–.004	.03	167.2	–0.14	1.0
Eye Contact – No looking at object vs. No Eye Contact – No looking at object	.006	.03	166.4	0.21	1.0

Notes. Results are based on the pairwise contrasts between the estimated marginal means of all conditions, inferred from the main GLMM fitted to the data by using the R-package "emmeans". P-values are adjusted via Bonferroni correction for six tests.

Table 4

Results from post-hoc one-sample t-tests testing the novelty preference score within the four conditions of Experiment 2 against chance

Condition	M*(SD)*	*t*	*df*	*p*	*d*
First-Party Eye Contact Looking at object	.61*(.13)*	4.62	31	<.001***	.82
No First-Party Eye Contact Looking at object	.49*(.11)*	–0.44	31	.66	–.07
First-Party Eye Contact No Looking at object	.48*(.12)*	–0.79	31	.43	–.14
No First-Party Eye Contact No Looking at Object	.50*(.12)*	0.12	31	.91	.02