BIROn - Birkbeck Institutional Research Online

ELSEVIER

Full length article

Probing communication-induced memory biases in preverbal infants: Two replication attempts of Yoon, Johnson and Csibra (2008)☆

Priya Silverstein[a],[*], Teodora Gliga[b],[c], Gert Westermann[a], Eugenio Parise[a]

[a] Department of Psychology, Lancaster University, Bailrigg, Lancaster, LA1 4YF, UK
[b] Centre for Brain and Cognitive Development, Department of Psychological Sciences, Birkbeck, University of London, London, WC1E 7HX, UK
[c] School of Psychology, University of East Anglia, Norwich Research Park, Norwich, Norfolk, NR4 7TJ, UK

ARTICLE INFO

ABSTRACT

In a seminal study, Yoon, Johnson and Csibra [PNAS, 105, 36 (2008)] showed that nine-month-old infants retained qualitatively different information about novel objects in communicative and non-communicative contexts. In a communicative context, the infants encoded the identity of novel objects at the expense of encoding their location, which was preferentially retained in non-communicative contexts. This result had not yet been replicated. Here we attempted two replications, while also including a measure of eye-tracking to obtain more detail of infants' attention allocation during stimulus presentation. Experiment 1 was designed following the methods described in the original paper. After discussion with one of the original authors, some key changes were made to the methodology in Experiment 2. Neither experiment replicated the results of the original study, with Bayes Factor Analysis suggesting moderate support for the null hypothesis. Both experiments found differential attention allocation in communicative and non-communicative contexts, with more looking to the face in communicative than non-communicative contexts, and more looking to the hand in non-communicative than communicative contexts. High and low level accounts of these attentional differences are discussed.

1. Introduction

Humans are expert learners. We learn implicitly, through mechanisms like statistical learning (Fiser & Aslin, 2001; Kirkham, Slemmer, & Johnson, 2002; Newport & Aslin, 2004), and explicitly from others through social learning (Csibra & Gergely, 2009, 2011; Tomasello, Carpenter, Call, Behne, & Moll, 2005). Social learning can occur either through observation (Meltzoff, 1988a, 1988b; Meltzoff & Moore, 1989), or through pedagogy, or explicit teaching (Csibra & Gergely, 2009; Csibra, 2007; Tomasello et al., 2005). Although teaching usually involves language, knowledge transfer can also occur in its absence. Two types of communicative cues have been suggested to be key to information transmission through teaching: ostensive cues such as direct eye contact or infant directed speech (IDS) convey the intention of communication, and referential cues such as pointing or gaze shifts direct attention to the source of the information to be learned (Csibra & Gergely, 2009; Csibra, 2010). According to the Natural Pedagogy theory (Csibra

& Gergely, 2009) ostensive communicative cues signal to infants when to learn culturally relevant kind-generalizable information about an object. In the presence of these cues, infants would be biased to encode surface features, which support learning about object kinds, over spatio-temporal information.

One method to investigate how infants encode object properties is the violation of expectation paradigm (VoE). The VoE paradigm is based on the assumption that infants look longer at events that violate their expectations (Onishi & Baillargeon, 2005; Teglas et al., 2011; Woodward, 1998), including when features of an object change (Krøjgaard, 2009; Mareschal & Johnson, 2003). Yoon, Johnson, and Csibra (2008) used a VoE paradigm to test the hypothesis that being communicated to should bias infants to encode surface features. In their study, the authors presented the infants with videos of communicative and non-communicative scenarios. The communicative videos included IDS, direct eye contact, and pointing, whereas the non-communicative videos included adult-directed speech (ADS), no direct eye contact, and reaching. In communicative scenes, an actress said 'Hey baby!' in IDS, while engaging in direct eye contact, and then pointed towards a novel object out of reach on the left or right side of the scene. In non-communicative scenes, an actress said 'What's this?' in ADS, while looking at the object, and then reached towards the object. Screens then occluded the object and actress. After a few seconds, the occluders opened to reveal the object again. At the point of reveal, either the identity or location of the object had changed, or no change occurred.

Infants looked longer at the identity change in the communicative condition and at the location change in the non-communicative condition (both in terms of first look and total look length). The authors concluded that infants encoded the identity of the object after being communicated to, as this was relevant to kind-generalizable learning. In contrast, infants encoded the location of the object in the non-communicative condition, due to this being the default, or perhaps the attempted reach enhancing the perceived graspability of the object. This double dissociation in the encoding of identity and location information suggested that the communicative cues did not merely increase overall memory, but elicited a specific memory bias towards identity information.

There have been few papers so far attempting to replicate or extend this finding. Okumura, Kobayashi, and Itakura (2016) found that in a live study, infants showed an identity bias in a direct gaze condition. However, the authors did not replicate the finding of a location bias for the condition with no direct gaze, instead finding encoding of both identity and location. The authors suggest that, due to the video deficit effect (Anderson & Pempek, 2005), infants may have performed better in their study than in the original study, therefore managing to encode both identity and location in the non-communicative condition. Their findings suggest that instead of identity being preferentially encoded by infants after viewing communicative scenes, it may be that infants are able to encode both spatiotemporal and recognition-relevant features, but ostensive signals disrupt location encoding. Two studies following up on this result in adults drew the same conclusions (Marno, Davelaar, & Csibra, 2014, 2016), finding encoding of both identity and location information in the non-communicative condition, and only encoding of identity in the communicative condition. As there are only two studies investigating communicatively induced memory biases in infants, we felt it was necessary to replicate the original finding before extending this research ourselves.

After being sent example videos from one of the original authors we noticed that in the communicative videos, the actress pointed for longer than she reached (6.8 s compared to 4.5 s). This difference raises the possibility that a longer duration of having the hand on screen could be responsible for inducing an identity memory bias. Therefore, in our stimuli both types of actions were performed twice for the same duration. Also, in the original study, the actress had bars stopping her from being able to reach the object. We did not use bars in our stimuli, instead having the actress be too far away from the object to reach it, in order to conceptually replicate the idea that the actress was unable to reach the objects, but without obscuring her. Lastly, we used eye tracking to investigate infants' distribution of attention while observing the communicative or non-communicative scenes. Although Yoon et al. (2008) compared overall looking to the action scenes and reported no difference between conditions, this does not speak to where the infants were looking while viewing these scenes. A difference in attention allocation in the two contexts could be responsible for the memory biases observed. For example, as preverbal infants tend to follow referential cues only when these are preceded by ostensive cues (Senju & Csibra, 2008; but see Gredebäck, Astor, & Fawcett, 2018; Szufnarowska, Rohlfing, Fawcett, & Gredebäck, 2014), we thought that perhaps in the communicative context infants would look more directly towards the objects, and that this could enhance the encoding of features. As this part of the study was exploratory, we did not pre-specify specific hypotheses relating to looking during the pre-occlusion section of the videos. We report two replication attempts. The first replication, Experiment 1 (bar the changes outlined above), followed the methodological details reported in Yoon et al. (2008). In the second replication, Experiment 2, we made slight changes to the method and exclusion criteria following guidelines provided by Csibra (personal communication, November 2017). Both replication attempts were pre-registered, and all data, code, supplementary results and materials are openly available on the Open Science Framework (OSF) (https://osf.io/77gpt/).

2. Experiment 1

2.1. Methods

Experiment 1 was conducted following the methods section of Yoon et al. (2008), bar some changes outlined in the introduction. Participants were recruited from a database of families at an infancy lab of a UK university, and were given a book as a gift for participation and £10 travel reimbursement. Parents gave informed, written consent before participation, and were free to withdraw their consent. All data were kept confidential. Both experiments were approved by the university ethics committee and adhered to the British Psychological Society guidelines.

2.1.1. Participants

2.1.1.1. Replication. Forty-two normally developing 9-month-old infants took part in the experiment. Of these twenty-four were included in the replication analysis (mean age: 274 days; range: 258 days to 286 days; 9 female; 23 Caucasian; 20 monolingual English). Exclusion criteria matched those used in Yoon et al. (2008): Infants were excluded for ceiling looking time for all trials (n = 4), experimenter error (n = 4), fussiness (n = 7), and not looking during one or more occlusion events (n = 3). An occlusion event in this experiment was defined as the time between the first frame of the clip when the occluder starts closing, and the first frame when the occluder is fully closed.

2.1.1.2. Scene analysis. We were able to use less stringent exclusion criteria for this analysis as our exclusion criterion of watching the entire occlusion event was irrelevant when looking at infant looking before the occlusion. Of the forty-two infants who took part in the experiment, 40 were included in the scene analysis (mean age: 276 days; range: 258 days to 322 days; 17 female; 39 Caucasian; 20 monolingual English). Two infants were excluded due to experimenter error.

2.1.2. Stimuli

We created the video stimuli and digitally added objects (on the left or right side of screen) and occluders on each clip. Infants first saw two familiarization trials. These familiarization videos were 29 s long and consisted of the actress moving around slightly to upbeat music, while looking either at the infant with direct gaze and smiling (communicative condition), or at the object with intrigue (non-communicative condition) (6 s). After this, yellow screens occluded the object and actress (3 s), there was a short break where the occluders stayed closed (5 s), the object screens revealed the object again (with no change) (2 s), and the object remained on screen for a maximum of 15 s (less if the infant looked away for two seconds, in which case the next trial was advanced). We always used the same two objects, but counterbalanced across participants for actress, side and order.

In the test videos (Fig. 1), there was first an introductory sentence produced by the actress ("Hey baby!" with direct gaze for the communicative videos, and "What's that?" with gaze to the object for the non-communicative videos) (4 s). This was followed by the action being executed once towards an object to the front of the actress on the left or right side (a point or a reach towards the object as if to try and grasp it) (4 s). The actions were completed at an equal distance away from the object, and were of the same duration. After completing the action once, the actress returned to the resting position and said either "Wow" (with direct gaze) while waving at the infant (communicative) or "Hmm" (without direct gaze) with a hand on her chin (non-communicative) (4 s). After this, she produced the pointing or reaching action again (4 s). Following this, screens moved to occlude both the object and the actress (2 s). The occluders stayed on the screen (5 s), after which the object screens reopened (2 s) to reveal the object. There had either been no change to the object, or it had changed in either identity or in location. The objects stayed on screen for a maximum of 15 s, or until infants looked away for 2 s. Test videos were 37 s long. Occluders produced sounds when opening and closing to direct infant attention to these events.

We obtained the objects from the Noun database (Horst & Hout, 2015), and chose 6 pairs with medium similarity ratings. We used the first of each of these pairs as the initial object, and the paired objects for the identity change condition (i.e., when the object changed identity, the second object in the pair was what it changed into). Every infant saw all 6 of the objects at test. Which object was shown for which condition combination was pseudo-randomised into 8 trial orders (with 3 infants viewing each order). Which actress played which role (communicative or non-communicative) was counterbalanced across infants. Object, side of action, condition and outcome were pseudo-randomised in the 8 possible trial orders. Additionally, infants never saw more than two trials similar on any factor (e.g. no more than two left actions) in a row. All videos are openly available on the OSF.

2.1.3. Procedure

Infants sat on their caregiver's lap during the experiment in front of a 23-inch screen (seated approximately 0.6 m away). An eye-tracker (Tobii X120) captured infant looking times and locations on screen for the action scene analysis. We used Tobii studio 3.3.1 to present stimuli and gather eye-tracking data. A camera placed above the screen fed input directly into Tobii studio, from which videos of infants were later exported for offline looking time coding. We performed a 9-point calibration for all infants before beginning the experiment. After this calibration, we instructed parents not to talk to or interact with their infant, and the experiment began.

Infants saw 2 familiarization videos (1 communicative, 1 non-communicative) followed by 6 test videos (for each of the communicative or non-comunicative familiarization video there was one no change, one identity change, and one location change test video). Therefore, each infant contributed one trial per sub-condition to the replication analysis, and three trials per condition to the scene analysis.

2.1.4. Analysis

2.1.4.1. Replication. We performed all analyses on raw data in R (R Core Team, 2017). All code is openly available on the OSF. Videos of infants were exported from Tobii studio and infant looking was blind coded in ELAN (2019)(Version 5.0.0). A second blind coder performed secondary coding on 20% of videos. We found a high degree of reliability between coders: the average measure ICC was 0.97 with a 95% confidence interval from 0.95 to 0.99. The length of the first look was defined as the time between the infant looking towards the screen and their first look away, beginning at the first frame of the occluders opening. The total looking time was defined as the cumulative length of time of all looks towards the screen, beginning at the first frame of the occluders opening, and ending at the first frame of the next attention getter.

We replicated the analysis by Yoon et al. (2008). We carried out a 2 × 3 repeated ANOVA on the length of first look to screen after object reveal as a function of action (communicative vs. non-communicative) and outcome (identity change, location change, no
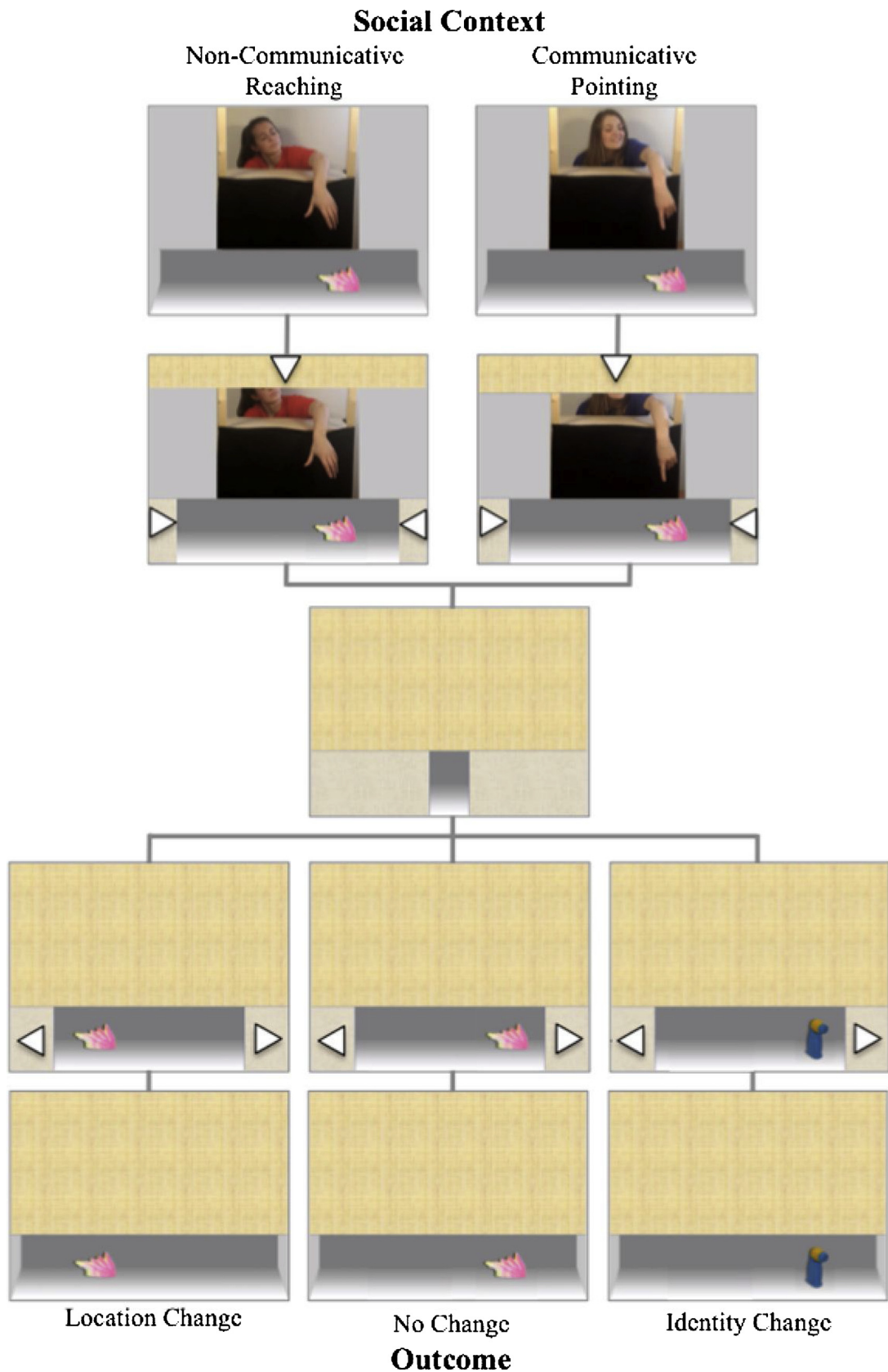
Fig. 1. Flow chart of stimuli in the test trials. Full videos are available on the OSF.
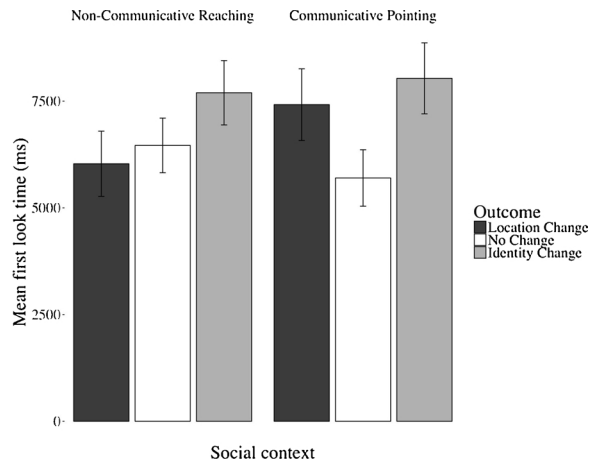
Fig. 2. Experiment 1: Mean duration of first look with SE bars for each outcome grouped by communicative context.

change), with follow up one-way ANOVAs and parametric and non-parametric pairwise tests. The same analyses were conducted for total looking times.

2.1.4.2. Scene analysis. We performed all analyses on raw data in R. All code is openly available on the OSF. We exported raw data from Tobii Studio 3.3.1, and visualized and analyzed them in R using the eyetrackingR package (Dink & Ferguson, 2016). First, Areas of Interest (AOIs) were created for face (630 × 310 pixels), hand (385 × 340 pixels), and object (380 × 340 pixels) areas of the videos. The data were then visualized as a timecourse. We performed a bootstrapped cluster based permutation analysis (Maris & Oostenveld, 2007) to establish during which time-points conditions differed significantly. This involved running a test on each time bin (17 ms) that quantified a significant difference between actions (communicative vs. non-communicative). We grouped into clusters the adjacent bins that showed a significant difference. We then shuffled the data and performed this same test on one thousand iterations of the shuffled data. This produced a table of the probability of each cluster appearing under the null hypothesis. Clusters that had a probability of less than 5% of appearing under the null hypothesis (i.e., p < 0.05) were considered to be significant. This test accounts for both Type 1 and Type 2 errors, by controlling the false-alarm rate while sacrificing little sensitivity (Maris & Oostenveld, 2007).

2.2. Results

2.2.1. Replication: first look length

We compared the length of first look in a 2 × 3 ANOVA with action (communicative vs. non-communicative) and outcome (location change, identity change, no change) as within-subject factors (Fig. 2). There was a significant main effect of outcome [$F(2,46) = 4.15$, $p = 0.02$, $\eta_p^2 = 0.15$]. There was no significant main effect of action [$F(1,23) = 0.2$, $p = 0.66$] or significant interaction between action and outcome [$F(2,46) = 1.32$, $p = 0.28$]. Paired comparisons by parametric (Student's t) and nonparametric (Wilcoxon rank-sum) tests were carried out to assess the effect of each outcome on the looking times. Infants looked longer to the identity change than the no change outcome, regardless of communicative context. Looking was significantly longer for the identity change than the no change outcome [$t(23) = 2.79$, $p = 0.008$, $\eta_p^2 = 0.26$; Wilcoxon's $Z = -2.45$, $p = 0.01$], whereas looking time for the identity change and location change [$t(23) = 1.47$, $p = 0.15$] and location change and no change [$t(23) = 0.95$, $p = 0.35$] outcomes did not differ significantly. A nonparametric McNemar test found that this difference could be attributed to some of the infants rather than the entire group. Six of the 24 infants showed the behaviour reported in Yoon et al.'s study (longer looking at identity change than no-change after communicative scenes and longer looking at location change than no-change after non-communicative scenes), 9 infants showed only the identity bias (longer looking at identity change than no-change after communicative scenes), 5 infants only showed the location bias (longer looking at location change than no change after non-communicative scenes), and 4 infants showed the opposite pattern in both contexts (longer looking at the identity change than no-change after non-communicative scenes and longer looking at location change than no change after communicative scenes). This distribution was not significantly different from chance (McNemar's $p = 0.42$).

2.2.2. Replication: total look length

Total looking length was compared in a 2 × 3 ANOVA with action (communicative or non-communicative) and outcome (location change, identity change, no change) as within-subject factors (Fig. 3). There was no main effect of outcome [$F(2,46) = 0.66$, $p = 0.52$] or action [$F(1,23) = 0.01$, $p = 0.93$], showing that overall infants did not look longer at test in either the communicative or non-communicative condition. There was a significant interaction between action and outcome [$F(2,46) = 3.5$, $p = 0.04$, $\eta_p^2 = 0.13$]. We carried out separate 3-level one-way repeated-measures ANOVAs followed by paired comparisons by parametric (Student's t) and nonparametric (Wilcoxon rank-sum) tests to assess the effect of each outcome on the looking times. In communicative context
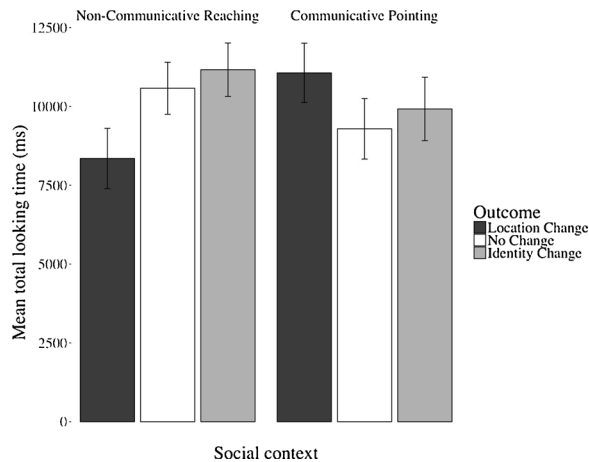
Fig. 3. Experiment 1: Mean duration of total look with SE bars for each outcome grouped by communicative context.

trials, there was no difference in looking time between the three outcomes [F(2,46) = 1.06, *p* = 0.35]. In non-communicative context trials, there was a main effect of outcome [F(2,46) = 3.91, *p* = 0.03, η_p^2 = 0.15]. Further analyses revealed significantly shorter looking for location change compared to both identity change [t(23) = 2.36 *p* = 0.03, ηp^2 = 0.19; Wilcoxon's Z = -2.34, *p* = 0.02] and no change [t(23) = 2.21 *p* = 0.04, η_p^2 = 0.18; Wilcoxon's Z = -1.76, *p* = 0.08] (note: when using a non-parametric test the difference between looking time to location change compared to no change was not significant). There was no difference between identity change and no change outcomes [t(23) = 0.61, *p* = 0.55].

2.2.3. Scene analysis

Figs. 3, 4 and S5 (see Supplementary materials on the OSF) show proportion looking to the face, hand, and object AOIs, respectively, for communicative and non-communicative scenes. Overall, infants showed similar looking patterns when viewing both types of scenes (looking towards the face the most, especially when the actress was speaking, looking towards the hand when the point/reach was being performed, and very little looking towards the object at any time point). A bootstrapped cluster-based permutation analysis found that during both of the periods where the hand action (point/reach) was not being performed (0–3000 ms, 6000–9000 ms), looking towards the face was significantly higher in the communicative condition (*p* = 0.027 and *p* = 0.048 respectively), and conversely, during the first time the action (point/reach) was performed (3000–6000 ms), looking towards the hand was significantly higher in the non-communicative condition (*p* = 0.024) (Fig. 4). Looking towards the object was very low overall and did not differ significantly between conditions. Overall looking time towards the scene did not differ between conditions [t(39) = 0.03, p = 0.98].

2.3. Discussion

2.3.1. Replication

Our results do not replicate the finding by Yoon et al. (2008) that infants show a memory bias for identity information after viewing communicative scenes, and a memory bias for location information after viewing non-communicative scenes. Neither do we find that communicative scenes disrupt the encoding of location information, which is preserved when viewing non-communicative scenes (Marno, Davelaar, & Csibra, 2016; Marno et al., 2014; Okumura et al., 2016). Instead, our results suggest that infants show longer looking to identity changes regardless of communicative context (measured by length of first look). However, we found that a minority of infants drove this effect, and that we do not find the results in the same direction for total look, weakening our belief that this truly indicates an identity memory bias. If we do interpret this result as an identity memory bias, this is surprising, given that the default for preverbal infants seems to be to encode location over surface features (Carey & Xu, 2001; Haun, Call, Janzen, & Levinson, 2006; Mareschal & Johnson, 2003; Xu & Carey, 1996). After this first replication attempt, we contacted Csibra, who generously provided detailed comments that led to some key methodological changes in Experiment 2.

2.3.2. Scene analysis

Our results showed that there are differences in where infants allocate their attention when viewing communicative and non-communicative scenes. Infants looked more towards the face when direct gaze and infant-directed speech were displayed, and more towards the hand when a reach was performed than when a point was performed. This shows that even before infants are actively communicating themselves, they are responding differently when they are being communicated to, compared to when they are not. As both communicative and non-communicative scenes involved speech and the same sequence of actions, we can reasonably assume that differences are due to more specific features of the two types of scene. This experiment cannot clarify whether these differences are due to low-level perceptual differences or a higher-level understanding of being communicated to.

We found that infants looked more towards the face when direct gaze and infant-directed speech were used. However, in these
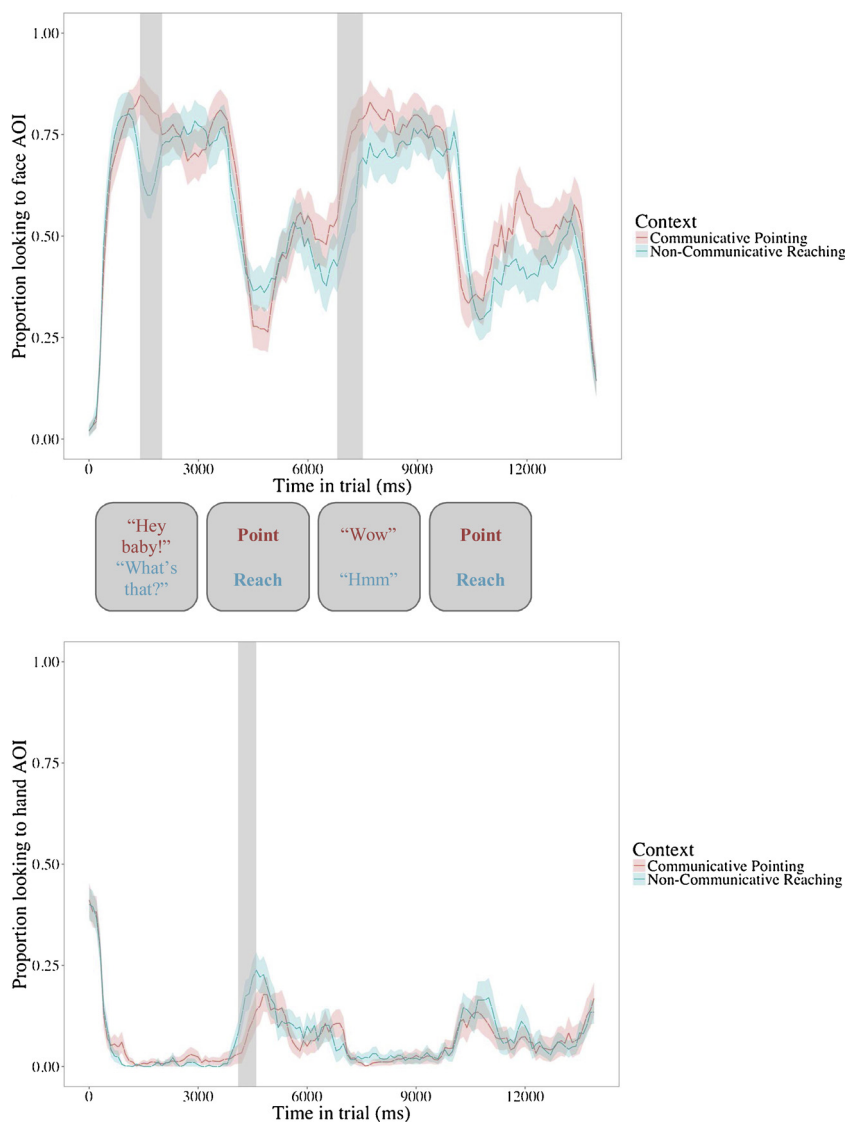
Fig. 4. Proportion looking to face (top) and hand (bottom) plotted over time (ms). N = 40.

scenes, this was also when the actress was waving to the infant. We could hypothesize that merely the movement of waving the hand is more salient than the moving of the head to look at the object in the non-communicative scenes, purely because there is more motion involved. Alternatively, consistent with Natural Pedagogy account (Csibra & Gergely, 2009), infants might look more towards the face when ostensive signals are present because they are prepared to learn from the interlocutor. Again, our results cannot differentiate between these two interpretations.

We also found that infants looked more towards the hand when it was reaching than when it was pointing. This difference in looking occurred at a different time-point to where we saw differences in looking towards the face, suggesting that this is not merely the other side of the coin (i.e. when infants are not looking at the face they are instead looking at the hand), and is in fact a different process at play. One low-level interpretation for this result could be that the hand occupies more space when it is a reaching hand than when it is a pointing hand, which could draw the infants' attention. Also, the reaching hand moves around a little, to show that the actress is unsuccessfully trying to reach the object, whereas the pointing hand does not move. Like the waving in the communicative videos, enhanced hand looking in this case could simply be the product of motion drawing the infants' attention. Alternatively, as the goal for the reach is to grasp the object, infants might fixate on the hand in order to see what happens next (i.e. whether the person manages to reach the object), whereas in the communicative condition, the goal of the point (to communicate) has been reached as soon as the infant perceives and understands it themselves. Again, these data do not speak to which of these interpretations is more likely. Our question was not the mechanisms behind any attentional differences, but instead whether attentional differences could be driving memory biases, and so further research should investigate the reasons for these differences. However, as we find an identity bias regardless of action condition, these differences in where infants allocate their attention cannot

be responsible for any memory biases in our experiment. It is also worth noting that infants are not looking at the object in either context, suggesting that they are not fully understanding the goal, as in both cases the goal is either to share attention about an object, or to reach the object. We know that 12-month-olds anticipate the goal of reaching actions by looking at the object, but 6-month-olds do not (Falck-Ytter, Gredebäck, & von Hofsten, 2006), so perhaps the infants in our experiment are too young to fully comprehend these goal directed actions (but see Kanakogi & Itakura, 2011).

3. Experiment 2

After communication with Gergely Csibra (personal communication, November 2017), we were made aware of some important methodological differences between our replication attempt and the original study. Most importantly, we found that our interpretation of 'occlusion event' was not the same as theirs. We had interpreted the 'occlusion event' as the time where the occluders were currently closing, whereas they had meant it to mean the time where the occluders were closing, *plus* the entire period where the object is currently occluded. As the original exclusion criteria specified that infants should be excluded if they had not watched all of the occlusion events without looking away, this meant that we had essentially used a different exclusion criterion. When we went back to our data to check which infants would still be included with this new, very strict criterion, we found that none of them would be. Further discussion with Csibra made it apparent that in Yoon et al. (2008) the occlusion time had been wrongly reported as five seconds, when in fact, in the original stimuli this was actually only three seconds. This longer occlusion time, and the fact that we did not include any music during the occlusion period, may have been responsible for none of our infants showing continuous looking at the screen during the whole 5 s occlusion event for all six trials. Therefore, in the second replication attempt, we shortened the occlusion time to 3 s, added music to the occlusion period, and changed our exclusion criterion to match that of the original paper. Additionally, this second replication could serve as a confirmation for our action scene findings, as this analysis was exploratory in the first attempt.

3.1. Methods

Methods remained overall the same as in Experiment 1, but with the changes to occlusion duration and exclusion criteria discussed with Csibra.

3.1.1. Participants
3.1.1.1. Replication. Seventy-nine typically developing 9-month-old infants took part in the experiment, and of these twenty-four were included in the replication analysis (mean age: 273 days; range: 261 days to 287 days; 13 female; 22 Caucasian; 23 monolingual English). Infants were excluded for falling asleep (n = 1), ceiling looking time for all trials (n = 7), fussiness (n = 14), experimenter error (n = 1), and not looking during one or more occlusion events (n = 32). An occlusion event in this experiment was defined as the time between the first frame of the occluder beginning to close and the first frame of the occluder being fully open (this was a crucial difference to Experiment 1).

3.1.1.2. Scene analysis. Of the seventy-nine infants who took part in the experiment, sixty were included in the scene analysis (mean age: 272 days; range: 260 days to 289 days; 31 female; 55 Caucasian; 57 monolingual English). Infants were excluded for falling asleep (n = 1), fussiness (n = 11), less than 40% good eye-tracking data (n = 3), and experimenter error (n = 1).

3.1.2. Stimuli
Stimuli were the same as in Experiment 1, except for three changes. First (outlined above), the objects were occluded for three seconds instead of five. Second, there was a larger gap (roughly 5 times wider) between occluders when objects were fully occluded. This was to ensure infants could see that the object had not moved from one side to the other (and thus, a location change would be genuinely surprising). In order to make such gap larger, the third change was that the objects were made slightly smaller, and moved further apart, which also made the object spacing more comparable to the original study.

3.1.3. Procedure
The procedure was identical to that of Experiment 1.

3.1.4. Analysis
3.1.4.1. Replication. The analyses were identical to those of Experiment 1. Secondary coding was performed on 20% of videos by a second blind coder. A high degree of reliability was found between coders. The average measure ICC was 0.95 with a 95% confidence interval from 0.90 to 0.98.

3.1.4.2. Scene analysis. The analysis was identical to that of Experiment 1.

3.2. Results

3.2.1. Replication: first look length
The 2 × 3 ANOVA revealed no significant main effect of outcome [$F(2,46) = 0.13$, $p = 0.88$] or action [$F(1,23) = 0.62$, $p =$
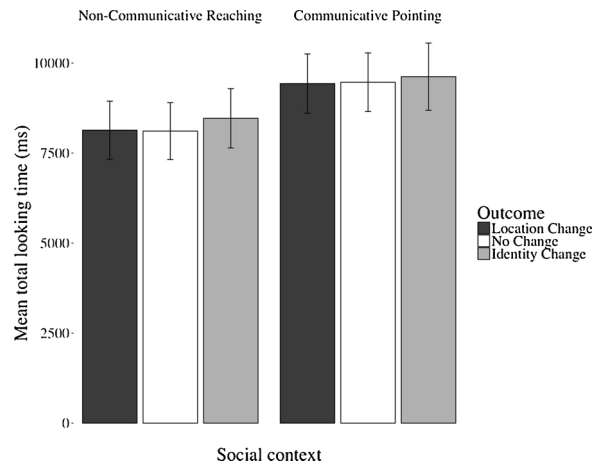
Fig. 5. Experiment 2: Mean duration of total look with SE bars for each outcome grouped by communicative context.

0.44] or significant interaction between outcome and action [F(2,46) = 1.32, p = 0.28] (Fig. S4) (Fig. 5). No statistical inference can be derived from this non-significant result (Lakens, McLatchie, Isager, Scheel, & Dienes, preprint). To determine whether the current data provides evidence for the null hypothesis (H0) relative to the alternative hypothesis (H1), a Bayes factor was conducted. Bayes factors (BF01) provide a measure of how likely the data are assuming H0 is true relative to how likely the data are assuming H1 is true. For the current analyses, a default Bayes factor with a wide cauchy distribution (scale of effect = 0.707) was calculated using the BayesFactor R package (Morey & Rouder, 2015), and yielded BF01 = 4.12. Thus, we can conclude that the data constitutes moderate evidence for the null hypothesis.

3.2.2. Replication: total look length

The 2 × 3 ANOVA revealed a significant main effect of action [F(1,23) = 4.33, p = 0.05, η_p^2 = 0.16], with infants looking longer to all outcome conditions after viewing the communicative videos. There was no main effect of outcome [F(2,46) = 0.09, p = 0.92], or interaction between action and outcome [F(2,46) = 0.01, p = 0.99] (Fig. S5 on OSF).

3.2.3. Scene analysis

Our findings from Experiment 1 for the scene analysis were exactly replicated (Fig. S6–S8 on OSF). During both of the periods where the hand action (point/reach) was not being performed, looking towards the face was significantly higher in the communicative condition (p = 0.002 and p = 0.025), and conversely, during the first time the action (point/reach) was performed, looking towards the hand was significantly higher in the non-communicative condition (p = 0.038). There were no significant differences in proportion looking to the object. Overall looking time towards the scene did not differ between conditions [t(59) = 0.86, p = 0.39].

3.3. Discussion

Our results show no evidence for memory biases, instead showing overall increased attention to the screen at test after viewing communicative videos. Our scene analysis results completely replicate findings from Experiment 1, suggesting that the attention allocation differences found are reliable.

4. General discussion

We have described two attempts to replicate the results of Yoon et al. (2008). In the original study, infants looked longer at an identity change following a communicative context, and longer at a location change following a non-communicative context. These results were interpreted as a preferential encoding of object identity in a communicative context. In our first replication attempt, we found longer looking at identity changes regardless of context. In our second attempt (which, following communication with Csibra (personal communication November 2017), was better matched to the original study in terms of stimuli and exclusion criteria), we found no memory biases at all, and instead just higher increased overall attention to the screen after viewing communicative videos compared to non-communicative videos. It is important to note the discrepancy in the results from Experiment 1 and Experiment 2. If we assume that the changes in stimuli and exclusion criteria in the two experiments did not have a meaningful impact on infant looking, these results may be due to random fluctuations due to small sample sizes, as when the two Experiments are combined, we get moderate to high support for the null hypothesis (see supplementary materials). Alternatively, if we assume that there were key differences in the stimuli for the two experiments, we should compare results from Experiment 2 to those found by Yoon et al. (2008), as these are better matched. These results replicate neither the original study showing a double dissociation of identity and location memory biases, nor do they replicate other findings of impaired location memory in communicative contexts (Marno et al., 2014, 2016; Okumura et al., 2016) (Table 1) or even any memory effects in general (Blaser & Káldy, 2010; Kibbe & Leslie, 2011). It may be

Table 1
Studies Investigating Communicatively Induced Memory Biases.

Study	N	Effect size	Non-communicative	Communicative
(Yoon et al., 2008 – infants)	24	$\eta p^2 = 0.41$	Location encoding	Identity encoding
(Okumura et al., 2016 – infants)	28	$\eta p^2 = 0.15$	Location & identity encoding	Identity encoding
(Marno et al., 2014 – adults)	24	$\eta p^2 = 0.16$	Location & identity encoding	Identity encoding
(Marno et al., 2016 – adults)	80	$\eta p^2 = 0.20$	Location & identity encoding	Identity encoding
The current paper: Experiment 1	24	$\eta p^2 = 0.15$	Identity encoding	Identity encoding
The current paper: Experiment 2	24	NA, but BF01 = 5	No encoding	No encoding

that the method itself is insensitive to measuring infant object memory in this specific paradigm.

Our results could differ from those found by Yoon et al. (2008) because of small differences in our stimuli such as the size of the faces, the use of superimposed objects as opposed to real ones, or the distance between the two objects (video examples from Yoon et al. (2008) are available on the OSF for comparison). It is also possible that some of these differences had an effect on the higher exclusion rate in our study compared to the original (70% vs. 57%) due to our videos being potentially less engaging. We did also purposefully make the methodological changes of the absence of the bars, and of the matching of the duration of actions in both conditions. However, we have no reason to assume that any of these factors would affect the postulated creation of memory biases through the presence or absence of communication. Despite this, it is still impossible to know whether these (or other) changes could be responsible for the difference in results. Regardless, these small changes are not accounted for in current theory, which would predict that with our setup we would find the same result as Yoon et al. (2008). The original study is a key piece of evidence for the claim of Natural Pedagogy that ostensive signals not only enhance attention, but also specifically induce an expectation to learn kind-generalizable information. If small differences to a paradigm can disrupt this, then this might question the generality of the claims made by this theory, as the changes that we made should not affect the hypothesized mechanism.

It is possible that the current study is a Type 2 error. This seems unlikely, given that we didn't replicate in either of the two attempts, and find moderate support for the null hypothesis using Bayes Factor Analysis (and moderate to high support for the null hypothesis if we combine the results from Experiments 1 and 2 – see supplementary materials on the OSF). However, it does remain possible that the true effect size is smaller than that observed by Yoon et al. (2008), and that a higher-powered study is needed in order to find an effect. Further studies on communicatively induced memory biases in infants could also investigate small methodological changes, in order to see under what specific scenarios the original finding holds and advance theory about what information infants encode in communicative and non-communicative contexts. However, due to the extremely high exclusion rate in the second experiment (70% of infants tested excluded from the replication analysis), this may be a very demanding and resource consuming challenge. As out of six experiments (Table 1) only one has shown a specific identity memory bias induced by communication (as opposed to the loss of location memory), we believe that within this paradigm, the evidence against the experimental hypothesis is stronger than evidence for it, and we must consider the possibility that the original result may be a Type 1 error. In order to further study this hypothesis in a way that doesn't require more than 50% data loss, we suggest the development of a new paradigm.

Our attention allocation results are robust, with Experiment 2 completely replicating the results from Experiment 1. We found that at certain time points infants looked more to the face in communicative contexts than in non-communicative contexts, and more to the hand when it was reaching than when it was pointing. These results could be due to low-level perceptual differences between the two types of scene (e.g. with infants allocating their attention to where there is more movement), or to a high-level mentalistic interpretation of why infants would pay more attention to these areas (awaiting communication from the face in the communicative context, and awaiting the outcome of the reach in the non-communicative context). We believe there should be caution in attributing rich interpretations to phenomena that could also be explained by lean, attention-based interpretations (Haith, 1998; Heyes, 2016; Newcombe, 2002). We originally wished to investigate infant attention allocation in order to relate it to the observed memory biases. As we do not observe any differential memory biases for different contexts, we cannot relate these attention allocation differences to memory. What we can say is that attention allocation differences do not have an impact on what information is encoded or retained in our studies. Nonetheless, we feel that, when possible, eye-tracking data should be used in looking time studies to rule out attention allocation differences driving effects.

Declaration of interest

None.

Acknowledgments

References

Anderson, D., & Pempek, T. (2005). Television and very young children. *American Behavioral Scientist, 48*(5), 505–522. https://doi.org/10.1177/0002764204271506.

Blaser, E., & Káldy, Z. (2010). Infants get five stars on iconic memory tests. *Psychological Science, 21*(11), 1643–1645. https://doi.org/10.1177/0956797610385358.

Carey, S., & Xu, F. (2001). Infants' knowledge of objects: Beyond object files and object tracking. *Cognition, 80*(1–2), 179–213. https://doi.org/10.1016/s0010-0277(00)00154-2.

Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & Language, 25*(2), 141–168. https://doi.org/10.1111/j.1468-0017.2009.01384.x.

Csibra, G. (2007). Teachers in the wild. *Trends in Cognitive Sciences, 11*(3), 95–96. https://doi.org/10.1016/j.tics.2006.12.001.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*(4), 148–153. https://doi.org/10.1016/j.tics.2009.01.005.

Csibra, G., & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences, 366*(1567), 1149–1157. https://doi.org/10.1098/rstb.2010.0319.

Dink, J., & Ferguson, B. (2016). *eyetrackingR. R package version 0.1.6.* < URL: . > http://www.eyetracking-R.com.

ELAN (Version 5.0.0-beta) [Computer software]. (2017, April 18). Nijmegen: Max Planck Institute for Psycholinguistics. Retrieved from https://tla.mpi.nl/tools/tla-tools/elan/.

Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature Neuroscience, 9*(7), 878–879. https://doi.org/10.1038/nn1729.

Fiser, J., & Aslin, R. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science, 12*(6), 499–504. https://doi.org/10.1111/1467-9280.00392.

Gredebäck, G., Astor, K., & Fawcett, C. (2018). Gaze following is not dependent on ostensive cues: A critical test of natural pedagogy. *Child Development, 89*(6), 2091–2098. https://doi.org/10.1111/cdev.13026.

Haith, M. (1998). Who put the cog in infant cognition? Is rich interpretation too costly? *Infant Behavior and Development, 21*(2), 167–179. https://doi.org/10.1016/s0163-6383(98)90001-7.

Haun, D., Call, J., Janzen, G., & Levinson, S. (2006). Evolutionary psychology of spatial representations in the hominidae. *Current Biology, 16*(17), 1736–1740. https://doi.org/10.1016/j.cub.2006.07.049.

Heyes, C. (2016). Born pupils? Natural pedagogy and cultural pedagogy. *Perspectives on Psychological Science, 11*(2), 280–295. https://doi.org/10.1177/1745691615621276.

Horst, J., & Hout, M. (2015). The Novel Object and Unusual Name (NOUN) database: A collection of Novel images for use in experimental research. *Behavior Research Methods, 48*(4), 1393–1409. https://doi.org/10.3758/s13428-015-0647-3.

Kanakogi, Y., & Itakura, S. (2011). Developmental correspondence between action prediction and motor ability in early infancy. *Nature Communications, 2*(1), https://doi.org/10.1038/ncomms1342.

Kibbe, M., & Leslie, A. (2011). What do infants remember when they forget? Location and identity in 6-month-olds' memory for objects. *Psychological Science, 22*(12), 1500–1505. https://doi.org/10.1177/0956797611420165.

Kirkham, N., Slemmer, J., & Johnson, S. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition, 83*(2), B35–B42. https://doi.org/10.1016/s0010-0277(02)00004-5.

Krøjgaard, P. (2009). The human ability to single out and track specific objects through space and time: Origin and application. In H. Høgh-Olesen, P. Bertelsen, & J. Tønnesvang (Eds.). *Human characteristics: Evolutionary perspectives on human mind and kind* (pp. 89–116). (1st ed.). Cambridge Scholars Publishing.

Lakens, D., McLatchie, N., Isager, P. M., Scheel, A. M., & Dienes, Z. (2018). *Improving inferences about null effects with Bayes factors and equivalence tests* (2018, March 7) https://doi.org/10.17605/OSF.IO/QTZWR.

Mareschal, D., & Johnson, M. (2003). The "what" and "where" of object representations in infancy. *Cognition, 88*(3), 259–276. https://doi.org/10.1016/s0010-0277(03)00039-8.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024.

Marno, H., Davelaar, E., & Csibra, G. (2014). Nonverbal communicative signals modulate attention to object properties. *Journal of Experimental Psychology: Human Perception and Performance, 40*(2), 752–762. https://doi.org/10.1037/a0035113.

Marno, H., Davelaar, E., & Csibra, G. (2016). An object memory bias induced by communicative reference. *Acta Psychologica, 163*, 88–96. https://doi.org/10.1016/j.actpsy.2015.11.005.

Meltzoff, A. (1988a). Infant imitation and memory: Nine-month-olds in immediate and deferred tests. *Child Development, 59*(1), 217. https://doi.org/10.2307/1130404.

Meltzoff, A. (1988b). Imitation of televised models by infants. *Child Development, 59*(5), 1221. https://doi.org/10.2307/1130485.

Meltzoff, A., & Moore, M. (1989). Imitation in newborn infants: Exploring the range of gestures imitated and the underlying mechanisms. *Developmental Psychology, 25*(6), 954–962. https://doi.org/10.1037//0012-1649.25.6.954.

Morey, R. D., & Rouder, J. N. (2015). *Bayes factor: Computation of Bayes factors for common designs. R package version 0.9.12-2.* https://CRAN.R-project.org/package = BayesFactor.

Newcombe, N. (2002). The nativist-empiricist controversy in the context of recent research on spatial and quantitative development. *Psychological Science, 13*(5), 395–401. https://doi.org/10.1111/1467-9280.00471.

Newport, E., & Aslin, R. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology, 48*(2), 127–162. https://doi.org/10.1016/s0010-0285(03)00128-2.

Okumura, Y., Kobayashi, T., & Itakura, S. (2016). Eye contact affects object representation in 9-month-old infants. *PLoS One, 11*(10), e0165145. https://doi.org/10.1371/journal.pone.0165145.

Onishi, K., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308*(5719), 255–258. https://doi.org/10.1126/science.1107621.

R Core Team (2017). *R: A language and environment for statistical computing.* URLVienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology, 18*(9), 668–671. https://doi.org/10.1016/j.cub.2008.03.059.

Szufnarowska, J., Rohlfing, K., Fawcett, C., & Gredebäck, G. (2014). Is ostension any more than attention? *Scientific Reports, 4*(1), https://doi.org/10.1038/srep05304.

Teglas, E., Vul, E., Girotto, V., Gonzalez, M., Tenenbaum, J., & Bonatti, L. (2011). Pure reasoning in 12-month-old infants as probabilistic inference. *Science, 332*(6033), 1054–1059. https://doi.org/10.1126/science.1196404.

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*(05), https://doi.org/10.1017/s0140525x05000129.

Woodward, A. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69*(1), 1–34. https://doi.org/10.1016/s0010-0277(98)00058-4.

Xu, F., & Carey, S. (1996). Infants' metaphysics: The case of numerical identity. *Cognitive Psychology, 30*(2), 111–153. https://doi.org/10.1006/cogp.1996.0005.

Yoon, J., Johnson, M., & Csibra, G. (2008). Communication-induced memory biases in preverbal infants. *Proceedings of the National Academy of Sciences of the United States of America, 105*(36), 13690–13695. https://doi.org/10.1073/pnas.0804388105.