



Proactive or reactive? Neural oscillatory insight into the leader–follower dynamics of early infant–caregiver interaction

Emily A. M. Phillips^{a,1} , Louise Goupil^b, Megan Whitehorn^a, Emma Bruce-Gardyne^a, Florian A. Csolsim^a, Ira Marriott-Haresign^a, and Sam V. Wass^a

Edited by Susan Goldin-Meadow, University of Chicago, Chicago, IL; received December 14, 2021; accepted November 29, 2022

We know that infants' ability to coordinate attention with others toward the end of the first year is fundamental to language acquisition and social cognition. Yet, we understand little about the neural and cognitive mechanisms driving infant attention in shared interaction: do infants play a proactive role in creating episodes of joint attention? Recording electroencephalography (EEG) from 12-mo-old infants while they engaged in table-top play with their caregiver, we examined the communicative behaviors and neural activity preceding and following infant- vs. adult-led joint attention. Infant-led episodes of joint attention appeared largely reactive: they were not associated with increased theta power, a neural marker of endogenously driven attention, and infants did not increase their ostensive signals before the initiation. Infants were, however, sensitive to whether their initiations were responded to. When caregivers joined their attentional focus, infants showed increased alpha suppression, a pattern of neural activity associated with predictive processing. Our results suggest that at 10 to 12 mo, infants are not routinely proactive in creating joint attention episodes yet. They do, however, anticipate behavioral contingency, a potentially foundational mechanism for the emergence of intentional communication.

infants | joint attention | dyadic interaction | intention | neural oscillations

Temporal and spatial coordination of one's gaze with another's, or joint attention, is fundamental to successful social interaction and shared cognition (1). Shared perception, afforded by joint attention, is thought to form the basis of shared intentions and human-specific forms of collective actions (2, 3). The ability to engage in reciprocally mediated joint attention, where both partners lead and follow each other's attention, develops toward the end of the first year and is a key milestone in developmental trajectories of language learning and social cognition (4–6). A distinction is made between “mutual” and “shared” joint attention. The former involves two individuals mutually attending to the same environmental stimulus together, at the same time; to be considered shared attention, however, mutual attention must be intentional—i.e., the partner who leads the other's attention toward a stimulus checks that the other partner has perceived it (e.g., by checking the partner's gaze), and the follower communicates that attention is shared (7).

The onset of intentional, proactive communication is debated (5), but a popular view has been that, already by the end of the first year, infants achieve episodes of joint attention through the establishment of shared intentionality; using ostensive signals deliberately, to direct and share the attention of a communicative partner (1, 5, 8). For example, 9- to 12-mo-old infants are thought to use declarative gestures and vocalizations to direct the attention of an experimenter (4) and modify their behavior depending on their success (9, 10). It is argued that infants' ability to proactively initiate and engage in triadic forms of shared attention toward the end of the first year is catalytic to early language acquisition and socio-cognitive learning, in creating a joint attentional frame where the focus and meaning of the adult partner's communication is not only shared between the adult and the infant but also common ground between them (i.e., both partners are attending to the same thing, and they both know that the other partner is attending to the same thing as them) (1, 7). More recently, it has been suggested that infants initiate joint attention not only to share attention but to directly elicit information from a social partner about their environment: communicating intentionally and actively to regulate when and how they learn (11). For example, infants aged 12 mo point in an interrogative manner (12, 13), and by 20 mo, they look toward their caregiver to ask for help when uncertain (14).

However, much previous work on joint attention development has been conducted using structured, experimental paradigms, where a researcher engages in clear, repetitive behaviors aimed at eliciting either a response to their initiations for shared attention in the infant or an initiation for shared attention by the infant. On each experimental trial, therefore, the adult's behavior is spatially precise and temporally stable (15): far from the fast-changing multi-layered complexity of naturalistic, free-flowing interactions (3, 16, 17). This saliency

Significance

Infants' ability to engage in joint attention predicts language development and socio-cognitive functioning. Despite its importance, we understand little about how joint attention is established. We use multi-level techniques to investigate whether infants deliberately create joint attention during naturalistic interactions. Our results suggest they do not: infants showed no evidence of social signaling before leading their partner's attention, and their endogenous oscillatory activity did not increase. Infants were, however, sensitive to their gaze being followed: when caregivers joined their attention, infants showed neural activity associated with anticipatory processing. Findings suggest infants do not actively control adults' attention—but perceive when adults respond to their initiations. Behavioral contingency may therefore be a key mechanism through which infants learn to communicate intentionally.

Author affiliations: ^aDepartment of Psychology, University of East London, London, UK E15 4LZ; and ^bCentre National de la Recherche Scientifique, Laboratoire de Psychologie et NeuroCognition, Université Grenoble Alpes, 38000 Grenoble, France

Author contributions: E.A.M.P. designed research; E.A.M.P., M.W., E.B.-G., F.A.C., and I.M.-H. performed research; E.A.M.P., S.V.W. and I.M.-H. analyzed data; S.V.W. secured funding; E.A.M.P., L.G., and S.V.W. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2023 the Author(s). Published by PNAS. This article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

¹To whom correspondence may be addressed. Email: u1920558@uel.ac.uk.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2122481120/-DCSupplemental>.

Published April 4, 2023.

and predictability may help infants deploy specific “communicative” behaviors at an age where they would not necessarily do so spontaneously during naturalistic interactions with their caregivers and without necessarily grasping the communicative nature of these behaviors yet. That is, over the course of the trial-by-trial repetitions, infants might learn that a change in their behavior affects change in the experimenter, giving rise to a behavioral response otherwise absent in sporadic, isolated, naturalistic contexts.

Consistent with this idea, recent micro-behavioral analysis of caregiver–infant tabletop play has shown that, in fact, during naturalistic interactions at the end of the first year, infants rarely engage in active attention-sharing behaviors. For instance, they have been found to look to their caregivers infrequently (17–19) and check the focus of their partners’ gaze before following their attention less than 10% of the time (15). Instead of routinely deploying communicative behaviors, infants at this age most often look directly toward objects, and join their caregivers’ gaze through attending toward the adults’ hands, as they manipulate the attended object (15, 19). Infrequent looks to the caregivers’ face by 13- to 14-mo-old infants has also been observed during free-moving, naturalistic interactions both in laboratory settings and home-based recordings (20, 21).

These findings from naturalistic interactions challenge the view that, toward the end of the first year, joint attention is already frequently achieved through proactive communication from the infant (i.e., using ones’ own gaze to signal communicative intention and using partner gaze to infer intention) and suggest that, at this point, shifts in infant attention might instead be mostly reactive to the behaviors of their partner. That infant attention is predominantly reactive during online social interaction has important implications for our current understanding of the learning mechanisms involved in the development of joint attention and how these mechanisms support language acquisition, as well as early socio-cognitive skills (18).

However, understanding how joint attention is established in caregiver–infant dyads, and addressing the mechanisms driving infant attention in shared interaction, is difficult using behavioral methods alone. This is because similar behaviors (e.g., looks toward objects or partners) can occur across different levels of attentional and intentional engagement (7, 17). Electroencephalography (EEG) provides a method to explore sub-second changes in neural activity at different oscillatory frequencies, which have previously been associated with broad mechanisms of cognitive engagement, in infancy and adulthood (16). Comparing EEG activity before, after, and during specific inter-dyadic moments in a free-flowing interaction thus allows insight into the fast-changing cognitive processes that govern how each partner’s attention is allocated.

Theta activity (3 to 6 Hz) is an oscillatory rhythm associated with endogenously driven attention and information-encoding processes in early infancy (22). In particular, EEG activity in the theta range has been found to increase over fronto-central electrodes during episodes of endogenously controlled attention. For example, theta activity increases where infants anticipate the next action of an experimenter (23), and while 10- to 12-mo-old infants view cartoon videos, theta activity increases over frontal electrodes during heart rate–defined periods of sustained attention (24). Both anticipatory looking and engaging in bouts of sustained attention rely on infants’ skill in endogenously controlling how their attention is allocated in the environment. Corresponding to this, fronto-central theta activity increases during self-guided object exploration, and theta activity occurring in the time before infants look toward an object has been found to predict the length of time infants pay attention to that object during solitary play (25, 26).

We hypothesized that, if controlled top-down processes drive infant attention when they lead their partner’s attention toward an object, theta activity would increase in the time window preceding infant-initiated looks to mutual attention compared to adult-led looks. To explore whether communicative signaling necessary for shared joint attention (7) also preceded moments of infant-led mutual attention, we compared the probability of infants looking to their partner or vocalizing in the time before look onset. Based on findings from experimental paradigms, an increase in ostensive signaling before infant-led attention was expected (9, 10). As a secondary research question, we also examined whether proactive engagement with their partner in the time before an infant-led look affected whether the look was followed by the caregiver. It was hypothesized that infant theta activity and their use of ostensive signals would increase in the time before infant-led looks to mutual attention compared to nonmutual attention.

A key process involved in the deliberate and intentional reorientation of a social partner’s attention in shared interaction is the anticipation of the partner’s response in the time after the initiation (7, 27–29). We therefore also compared infant neural oscillatory activity and ostensive signaling occurring immediately after the onset of infant- and adult-led looks to mutual attention. Naturalistic, observational studies have shown that infants are sensitive to the contingency of an adult partner. For example, responding contingently to an infant’s gestures immediately improves the quality and quantity of the attention that they pay to objects (30, 31), and when caregivers behave redirectively (i.e., non-contingently), infants’ visual attention durations immediately decrease (32, 33). To our knowledge, however, no previous work has investigated whether infants proactively anticipate, or predict, a response by the partner to their behavior, i.e., do they check whether their partner has perceived their new attentional focus, and communicate about it, once attention is shared (7).

As well as examining infants’ behavioral cues signaling the anticipation of joint attention after leading their partner’s attention, here, we also investigate whether we can identify neural markers of predictive processing in the time following gaze onset. In adults, alpha desynchronization is thought to represent release from inhibition during sensory information processing (34). Reduced alpha activity has been identified at the onset of a predicted stimulus (35), and in social paradigms, predicting the outcome of another person’s action is associated with alpha desynchronization over pre-central motor cortices (36–38). In infancy, similar patterns of alpha suppression (6 to 9 Hz) have been shown over motor areas when observing the predicted outcome of another individual’s manual behavior (39–41), and one recent study also showed alpha desynchronization over central–parietal areas when infants viewed the behavioral response of a video-recorded experimenter to their own behavior, who followed the infant’s gaze toward an object (42) (see also ref. 43). If infants anticipate the behavioral response of their partner where they lead a look toward an object, alpha desynchronization would be expected to occur in the time after infant-led looks to mutual attention, with infants encoding the predicted outcome of their initiation toward an object, on their partner’s behavior (i.e., following their attention).

Based on the view that infants deliberately and proactively initiate shared attention with their partner during social interaction (1, 8), we hypothesized that infant looks to their partner’s face would increase in the time after infant-led looks to mutual attention (i.e., that they would check whether their partner had followed their attention toward a new object of interest). It was further hypothesized that infant vocalizations would show some increase in the time after infant- and adult-led looks to mutual attention, with the infant communicating, intentionally, to their partner about the shared focus of attention (7, 8). Consistent with previous neurophysiological

findings (39–43), we hypothesized that if infants anticipate the behavioral contingency of their adult partner where they lead attention toward an object, decreased oscillatory activity in the alpha range (6 to 9 Hz) would occur in the time after infant-initiated looks to mutual attention, compared to adult-initiated looks, and infant-initiated looks to nonmutual attention.

Results

The Results section is in three parts. *Section 1—Descriptive Statistics* presents descriptive statistics on infant and adult gaze and vocal behavior. *Section 2—Before Look Onset: Are Infants Proactively Initiating Joint Attention Episodes?* compares the attentional, behavioral, and neural dynamics preceding a) infant-led vs. adult-led looks to mutual attention and b) infant-led looks to mutual vs. nonmutual attention. *Section 3—After Look Onset: Do Infants Anticipate Their Gaze Being Followed?* repeats this analysis in the time period following look onset. As behavioral cues are slower changing in comparison to EEG activity, a 5,000-ms time window was used to compare infant behavior in the time before and after a look onset (customary for this type of research; see ref. 15), while a 2,000-ms time window examined infant EEG activity. See Table 1 (*Materials and Methods*) for a detailed description of how adult-led looks to mutual attention and infant-led looks to mutual attention and nonmutual attention were defined.

Section 1—Descriptive Statistics. Prior to testing our main hypotheses, we conducted three descriptive analyses. First, before interpolating through looks to partner, we investigated the proportion of time that caregivers and infants spent vocalizing, looking to their partner, attending to objects, and inattentive during the interaction (Fig. 1 *A* and *B*). Second, after interpolating through looks to partner, we tested how many times per minute infants and adults engaged in episodes of mutual attention (infant or adult led) and nonmutual attention. Finally, we examined the length of infant

attention episodes and of caregiver–infant mutual attention episodes (Fig. 1 *C* and *D*).

Infants spent the majority of the time looking toward objects, whereas caregivers divided their attention between their infant and the objects (consistent with ref. 15; Fig. 1*A*). Infant vocalizations were infrequent, whereas adult vocalizations were more frequent (Fig. 2*B*). Comparisons between caregivers and infants were significant using two-tailed independent *t* tests: looks to objects [$t(72) = 10.81$, $P < 0.001$, $d = 0.84$], looks to partner [$t(72) = -14.01$, $P < 0.001$, $d = -1.02$], and vocalizations [$t(36) = -5.61$, $P < 0.001$, $d = -0.60$]. The proportion of time spent in states of inattention did not differ [$t(72) = 1.32$, $P = 0.198$, $d = 0.08$].

Infant object look durations were positively skewed before log transform (consistent with ref. 17), as were episodes of caregiver–infant mutual attention (Fig. 1*C*). The number of times each type of attention episode occurred per minute was similar for caregivers and infants, with leader looks to mutual attention the most infrequently occurring category (Fig. 1*D*). Two-tailed independent *t* tests showed that infants followed their partner's attention significantly more often per minute compared to their caregivers [$t(72) = 2.94$, $P = 0.004$, $d = 0.77$]; all other comparisons were not significant (leader to nonmutual looks [$t(72) = 1.49$, $P = 0.139$, $d = 0.35$]; leader to mutual looks [$t(72) = -0.73$, $P = 0.467$, $d = -0.14$]).

Section 2—Before Look Onset: Are Infants Proactively Initiating Joint Attention Episodes? This section is in two parts. *Infant-led vs. adult-led mutual attention* compares infants' use of ostensive signals and their neural oscillatory activity occurring before infant-led looks to mutual attention, and adult-led looks to mutual attention, in order to test for differences between infant- and adult-initiated mutual attention episodes. *Followed vs. not followed infant-led looks* subsequently compares infant-led object looks resulting in mutual and nonmutual attention in order to test for differences between infant-led looks that were followed, or not followed, by their adult partner.

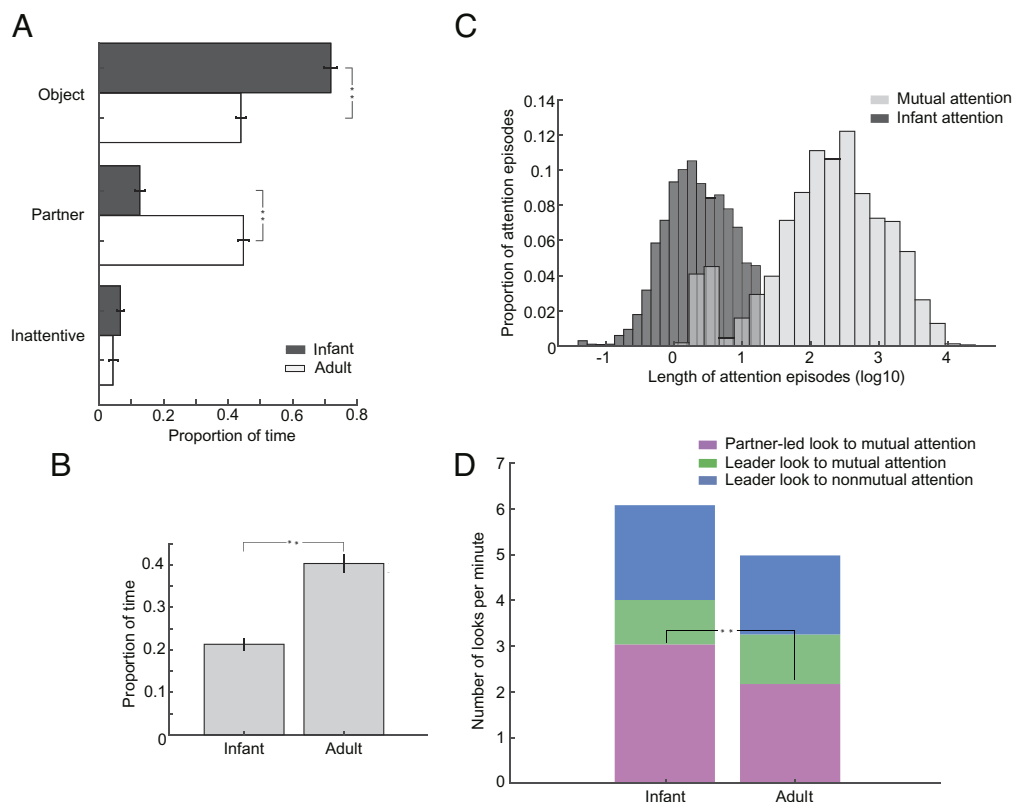


Fig. 1. Caregiver and infant attention and vocal behavior in shared play. (*A*) Bar plots show the proportion of time caregivers and infants spent looking to their partner, toward objects, and inattentive during the interaction. Two-tailed independent *t* tests ($n = 37$) compared proportions between caregivers and infants for each look category ($*P < 0.05$, $**P < 0.01$; error bars show the SEM). (*B*) Bar plot shows the mean proportion of time caregivers and infants spent vocalizing during the interaction. A two-tailed independent *t* test ($n = 37$) compared proportions between caregivers and infants ($*P < 0.05$, $**P < 0.01$; error bars show the SEM). (*C*) Histograms show log-transformed infant object look durations and length of mutual attention episodes, across all types of looks, after interpolation. (*D*) Bar plots show the number of times infants and adults engaged in one of three possible attentional states per minute: partner-led looks to mutual attention, leader looks to mutual attention, and leader looks to nonmutual attention. Two-tailed independent *t* tests compared the number of attention episodes per minute between caregivers and infants for each look category ($n = 37$; $*P < 0.05$, $**P < 0.01$).

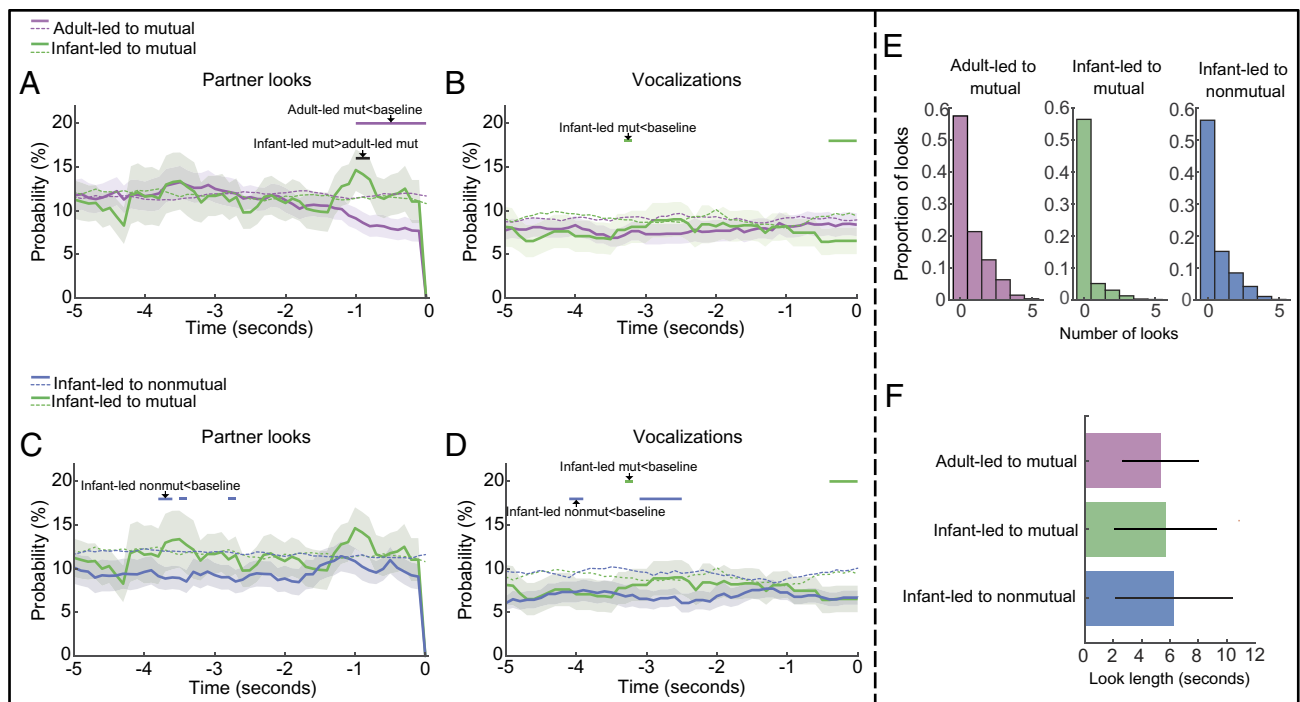


Fig. 2. Probability of ostensive signals (partner looks and vocalizations) and infant attentiveness in the time period before infant look onset. Probability time course for (A) partner looks before infant-led mutual attention vs. adult-led mutual attention, (B) vocalizations before infant-led mutual attention vs. adult-led mutual attention, (C) partner looks before infant-led mutual attention vs. infant-led nonmutual attention, and (D) vocalizations before infant-led mutual attention vs. infant-led nonmutual attention. In each case, shaded areas show the SEM [$n = 37$ for (A) and (C), $n = 19$ for (B) and (D)], and horizontal black lines show the areas of significant difference, between attention episodes, identified by the cluster-based permutation analysis (Monte Carlo P value < 0.05). Dotted lines show the baseline time series, plotted for each attention episode. Horizontal colored lines show the areas of significant difference between each attention episode and baseline, identified by the cluster-based permutation analysis (Monte Carlo P value < 0.05). (E) Histograms show the distribution of the number of object looks in the 5-s time period before look onset for each attention episode. (F) Bar plots show the mean length of infant attention toward the object immediately preceding look onset for all three attentional states. Error bars indicate the SEM ($n = 37$). Two-tailed paired t tests compared the length of infant attention toward the previous object, between each attentional state, which indicated no significant differences.

Infant-Led vs. Adult-Led Mutual Attention.

Ostensive signals and infant attention. First, we tested whether infants were more likely to use ostensive signals before an infant-led mutual attention episode compared to where they followed their caregiver's look into mutual attention. If true, this would support the hypothesis that infants proactively lead their caregiver's attention to objects during naturalistic tabletop play. To investigate this, we conducted a probability analysis examining ostensive signals in the time window $\pm 5,000$ ms relative to each look type.

For each look type (infant led to mutual and adult led to mutual), the frame at which an object look onset occurred was identified in the vocalization and partner look time series separately, and the 5,000 ms preceding look onset extracted. The probability of the behavior occurring at each 20-ms frame was then calculated as the proportion of looks where each ostensive signal (looks to the partner's face and vocalizations) was present in that frame. Results of the probability analysis are presented in Fig. 2 A–D. Cluster-based permutation analysis (*Materials and Methods*) indicated that infants were significantly more likely to look toward their caregiver in the time period immediately preceding an episode of infant-led mutual attention compared to an adult-led mutual attention episode (Fig. 2A). There were no other significant differences between look types (Fig. 2 A and B).

Any significant difference between adult-led and infant-led attention could, however, be driven either by an increase relative to baseline in looking prior to infant-led attention episodes or by a decrease relative to baseline prior to adult-led attention episodes. To differentiate between these hypotheses, we generated a random probability time series of partner looks and vocalizations in the time immediately before and after each type of infant object look

by inserting a random event into each ostensive cue time series and extracting the 5,000 ms preceding the event. A cluster-based permutation analysis was again conducted using paired t tests to investigate where the behavioral time series differed from chance (Fig. 2 A–D). Results indicated that the probability of infants looking to their partner was below levels expected by chance in the 1-s time period before the onset of adult-led attention. Overall, then, these results suggest that infants are less likely to look to their partner during the time window preceding adult-led looks to mutual attention.

In addition, to investigate whether infant attentiveness differed in the time before look onset between each look type, using the uninterpolated gaze time series, we also examined how frequently infant attention changed in the 5,000-ms time period leading up to each attention episode and the length of infant gaze toward the previous object for each type of look (Fig. 2 E and F). Two-tailed paired t tests showed no difference in the number of object looks occurring 5 s before the onset of an infant-led look to mutual attention [mean = 1.85, SEM = 0.08] compared to an adult-led look to mutual attention [mean = 1.17, SEM = 0.06; $t(36) = -0.27$, $P = 0.792$, $d = -0.24$]; nor was there a difference in the length of infant attention to the previous object [$t(36) = -0.61$, $P = 0.544$, $d = -0.10$].

Neural oscillatory activity. We compared how the neural oscillatory activity differed in the time before infant-led and adult-led mutual attention episodes. Results of the time–frequency analysis are presented in Fig. 3. Two-dimensional cluster-based permutation analysis revealed no significant clusters of time*frequency points comparing between infant-led looks to mutual attention and adult-led looks to mutual attention. Three-dimensional cluster-

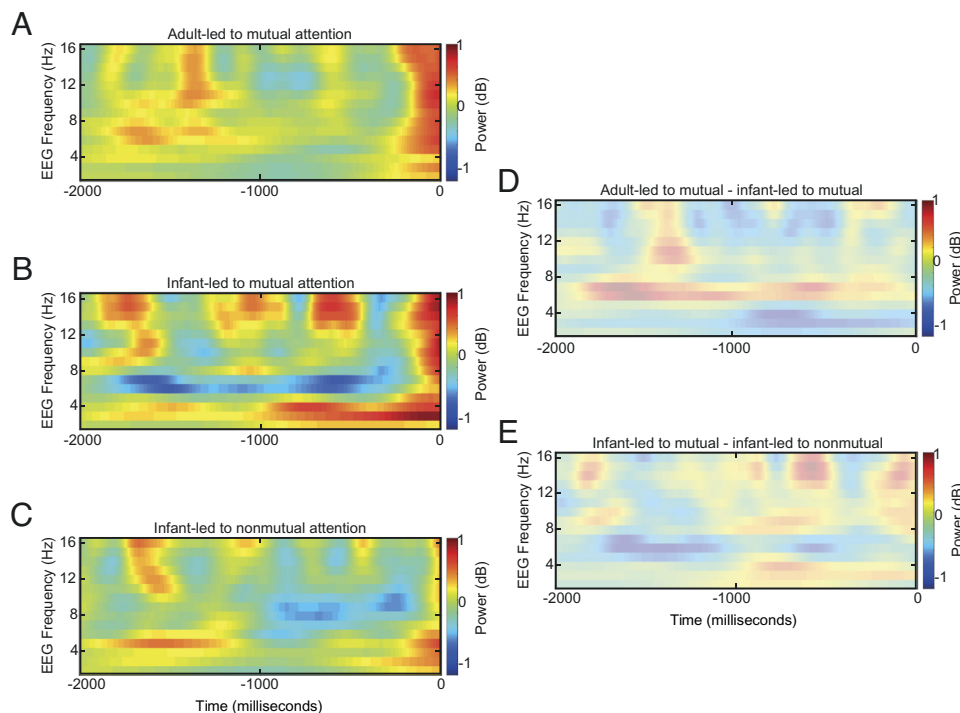


Fig. 3. Comparison of infant EEG activity in the 2,000 ms preceding infant- and adult-led attention episodes. Time–frequency plots show infant EEG activity (2 to 16 Hz) occurring 2,000 ms before look onset for (A) adult-led looks to mutual attention, (B) infant-led looks to mutual attention, and (C) infant-led looks to nonmutual attention over fronto-central electrodes (AF3, AF4, FC1, FC2, F3, F4, and Fz). Time 0 indicates infant gaze onset. (D) Difference in EEG activity between infant- and adult-led looks to mutual attention (adult-led – infant-led). Cluster-based permutation analyses showed no significant clusters of time*frequency points of the difference between attention episodes, and so no significant clusters have been highlighted. (E) Difference in EEG activity between infant-led looks to mutual and nonmutual attention (infant-led to mutual – infant-led to nonmutual). Cluster-based permutation analyses showed no significant clusters of time*frequency points of the difference between attention episodes, and so no significant clusters have been highlighted.

based permutation analysis, including all electrodes by time by frequency points, also revealed no significant clusters. Contrary to what would be expected if infants were deliberately orienting their partners toward objects when shifting their gaze to an unattended object, this primary analysis suggests that there were no significant differences in infants' neural activity in the time windows before they led their partner's attention toward an object.

In our naturalistic data, some of the epochs included in each look category will also have contained additional object and partner looks during the 2,000 ms before the onset of the look to which the data were event-locked. Even though eye movement-related artifacts were removed through ICA decomposition during preprocessing (*Materials and Methods*), we also conducted an additional analysis to examine the possibility that this may have contributed to the null result. The results suggested that it did not: the average proportion of looks with object and partner looks occurring in the time before look onset did not differ between attention episodes (*SI Appendix, Figs. S2–S4*). Conducting analyses including looks with no shifts in infant attention before look onset only was not possible due to low trial numbers (*SI Appendix, Fig. S1*). We therefore conducted a secondary analysis excluding neural activity within each look epoch, in the time before look onset, where an infant was not continuously focused on one object/ the partner (*SI Appendix, Fig. S5*). Two-dimensional cluster-based permutation analysis again revealed no significant clusters of time*frequency points (*SI Appendix, Fig. S5*).

For comparison with the behavioral analysis presented in *Infant-led vs. adult-led mutual attention*, *SI Appendix, Fig. S6* shows EEG activity over the same time windows (–5,000 ms; *SI Appendix, Fig. S6*). Two-dimensional cluster-based permutation analysis again revealed no significant clusters of time*frequency points, comparing between infant-led looks and adult-led looks to mutual attention.

Followed vs. not followed infant-led looks.

Ostensive signals and infant attention. Here, we also tested whether infants were more likely to use ostensive signals before an infant-led mutual attention episode, compared to an infant-led nonmutual

attention episode, in order to examine differences between infant-led looks that were followed, or not, by their adult partner. No significant differences were observed either in the likelihood of the infant looking to their caregiver in the time period preceding a look (Fig. 2C), or in the likelihood of the infant vocalizing (Fig. 2D). In addition, no significant differences were observed in the duration [$t(36) = 1.01$, $P = 0.321$, $d = 0.12$; Fig. 2F], or number [$t(36) = 1.45$, $P = 0.157$, $d = 0.24$] of infant objects looks in the time period preceding infant-led looks to mutual attention [mean = 1.85, SEM = 0.08], and nonmutual attention [mean = 1.06, SEM = 0.07; Fig. 2E].

Neural oscillatory activity. We also compared how neural oscillatory activity differed in the time before infant-led mutual attention episodes and infant-led nonmutual attention episodes (Fig. 3). No significant differences were observed using either the two-dimensional (Fig. 3) or the three-dimensional cluster-based permutation analyses. Again, the number of looks including object and partner looks before each attention episode did not differ (*SI Appendix, Figs. S2–S4*). EEG activity occurring 5,000 ms before look onset for each type of look is presented in *SI Appendix, Fig. S6*. Two-dimensional cluster-based permutation analysis revealed no significant clusters of time*frequency points of the difference between attention episodes.

Summary. In summary, these results suggest that there is little change in infants' behaviorally ostensive signaling before infant-led mutual attention episodes, compared with adult-led mutual attention (*Ostensive signals and infant attention*). The main finding was a decrease in infant looks to their caregiver in the time before adult-initiated mutual attention. There were no differences in infants' ostensive signaling between infant-led looks that were followed vs. not followed by their adult partner (*Ostensive signals and infant attention*).

The neural analyses suggested that there were no differences in neural oscillatory activity before infant-initiated and adult-initiated mutual attention (*Neural oscillatory activity*). There were also no differences in neural oscillatory activity between followed vs. not followed infant-led attention episodes (*Neural oscillatory activity*). There was thus very little evidence that 12-mo-old infants

proactively initiate joint attention with their partner during shared play.

Section 3—After Look Onset: Do Infants Anticipate Their Gaze Being Followed? In this section, we present a similar analysis to *Section 2—Before Look Onset: Are Infants Proactively Initiating Joint Attention Episodes?*, investigating change in infant behavior and neural oscillatory activity in the time period after look onset. Again, the section is organized in two parts: first, we examine mutual attention, comparing infant-led and adult-led mutual attention episodes (*Infant-led vs. adult-led mutual attention*). Second, we compare infant-led attention that was followed vs. not followed by their adult partner (*Followed vs. not followed infant-led attention*).

Infant-led vs. adult-led mutual attention.

Ostensive signals and infant attention. First, we tested whether infants were more likely to use ostensive signals during the time period after the start of infant-led, compared to adult-led mutual attention. To investigate this, we conducted the same probability analysis described in *Section 2*, extracting the 5,000 ms following

look onset from the vocalization and partner look time series. No significant difference in the likelihood of partner looks was observed, but a significant increase in the likelihood of infant vocalizations following adult-led mutual attention was shown (Fig. 4B). Baseline comparisons suggested that infant vocalizations significantly decreased from baseline in the time after infant-led looks to mutual attention, potentially driving this difference (Fig. 4B).

We also examined whether infant-led mutual attention episodes tended to be longer lasting than adult-led mutual attention (Fig. 4E). No significant difference was observed [$t(36) = -1.17$, $P = 0.248$, $d = -0.19$]. Finally, we examined the time interval it took caregivers to follow their infant's attention during infant-led looks to mutual attention (Fig. 4F). This analysis suggested that most looks were followed within 1 to 2 s after look onset [mean = 1.49 s, SEM = 6.91].

Neural oscillatory activity. In this section, we compare differences in infant EEG activity occurring over fronto-central electrodes after look onset for infant-led and adult-led looks to mutual attention (Fig. 5). Consistent with our hypothesis, infant-led mutual attention episodes led to a decrease in EEG power, particularly in the theta/

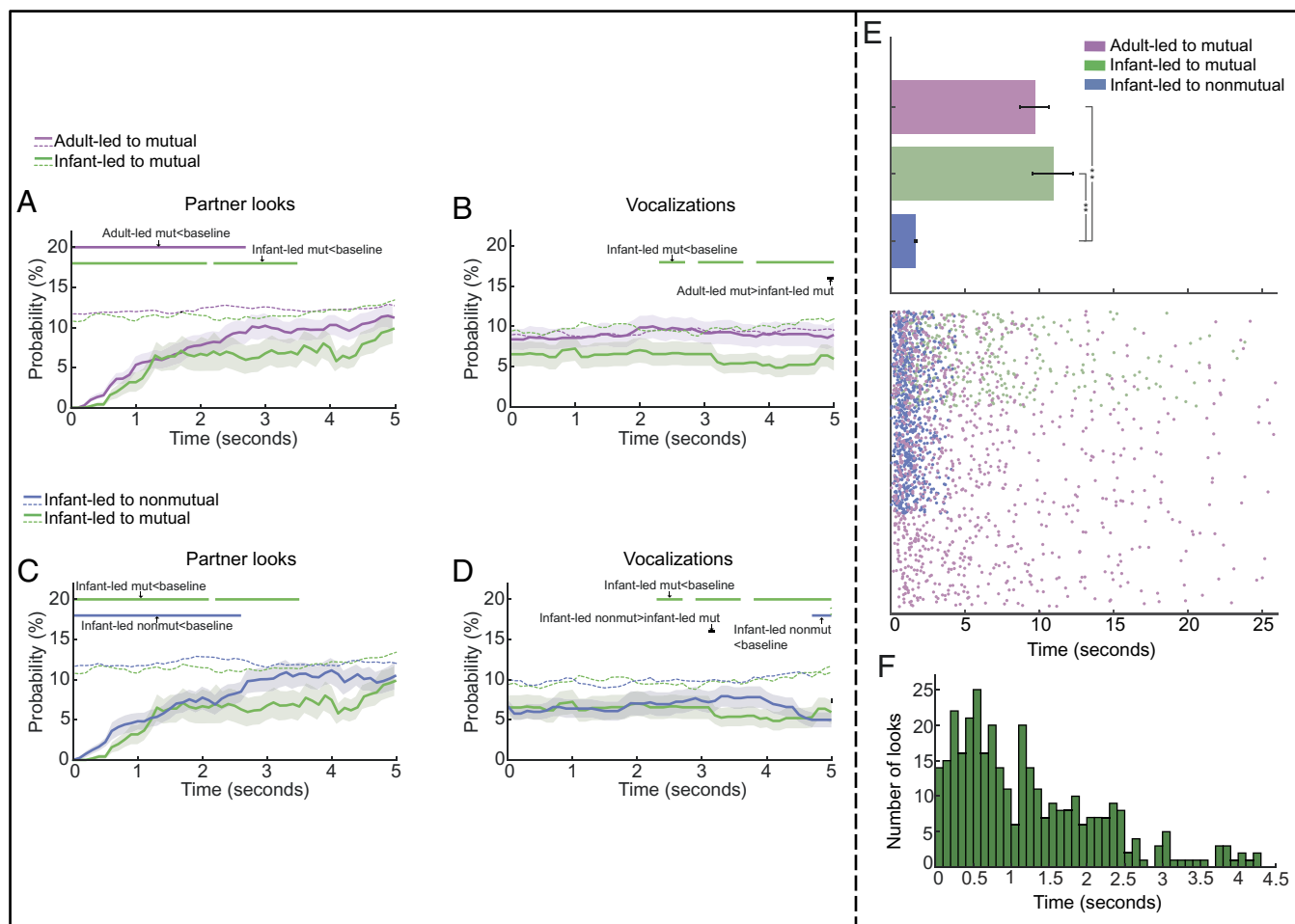


Fig. 4. Probability of ostensive signals (partner looks and vocalizations), infant attentiveness, and the time adults took to follow infant attention in the time period after infant look onset. Probability time course for (A) partner looks after infant-led mutual attention vs. adult-led mutual attention, (B) vocalizations after infant-led mutual attention vs. adult-led mutual attention, (C) partner looks after infant-led mutual attention vs. infant-led nonmutual attention, and (D) vocalizations after infant-led mutual attention vs. infant-led nonmutual attention. In each case, shaded areas show the SEM [$n = 37$ for (A) and (C), $n = 19$ for (B) and (D)], and horizontal black lines show the areas of significant difference, between attention episodes, identified by the cluster-based permutation analysis (Monte Carlo P value < 0.05). Dotted lines show the baseline time series plotted for each attention episode. Horizontal colored lines show the areas of significant difference between each look type and baseline identified by the cluster-based permutation analysis (Monte Carlo P value < 0.05). (E) Length of infant attention toward an object after look onset for each type of attention episode. The bar plot shows the mean length of infant attention averaged over participants [error bars show the SEM ($n = 37$)]; scatterplot shows the length of each individual look contributing to each object look category, across all participants, after outlier removal. Two-tailed paired t tests ($n = 37$) compared the difference in the length of infant attention between each type of attention episode (* $P < 0.05$, ** $P < 0.01$). (F) Histogram shows the distribution of the time it took caregivers to follow infant-led looks to mutual attention, across all looks, for all participants after outlier removal.

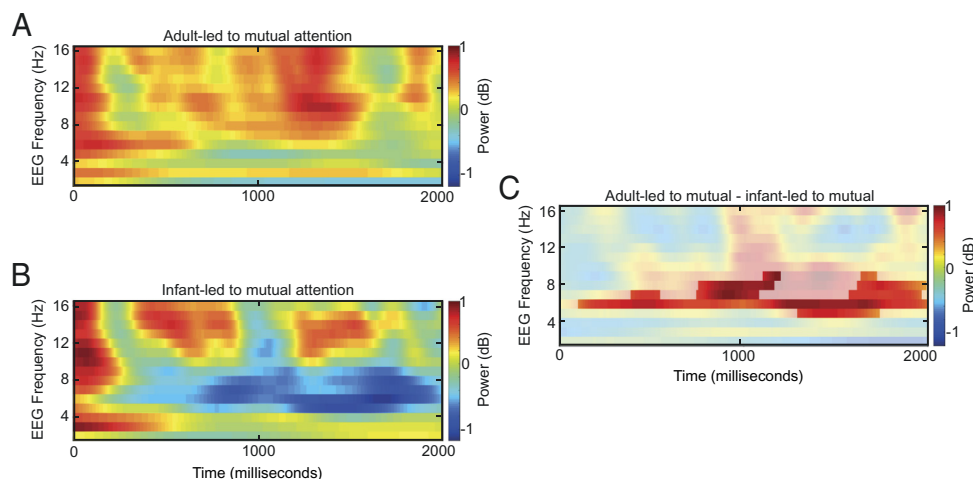


Fig. 5. Comparison of infant EEG activity in the 2,000 ms following infant- and adult-led attention episodes. Time-frequency plots show infant EEG activity (2 to 16 Hz) occurring 2,000 ms after look onset for (A) adult-led looks to mutual attention and (B) infant-led looks to mutual attention over fronto-central electrodes (AF3, AF4, FC1, FC2, F3, F4, and Fz). Time 0 indicates infant gaze onset. (C) Difference in EEG activity between infant- and adult-led looks to mutual attention (adult-led – infant-led); highlighted area shows the significant positive cluster identified by the cluster-based permutation analysis ($P = 0.003$). The cluster ranges from 5 to 9 Hz and from 92 to 2,000 ms after look onset.

alpha range toward the end of the 2,000-ms time period, as compared to adult-led looks (Fig. 5 *A* and *B*). The two-dimensional cluster-based permutation analysis identified a significant positive cluster with an average frequency of 7 Hz (ranging 5 to 9 Hz), 92 to 2,000 ms after look onset ($P = 0.003$; Fig. 5C). Three-dimensional cluster-based permutation analysis also revealed one trend-level positive cluster, with a wide topographical distribution, in the 5 to 9 Hz range ($P = 0.099$; *SI Appendix, Fig. S7*).

Again, due to the naturalistic nature of our data, some of the epochs included in these analyses contain additional object and partner looks. Similar to the pre-look analysis, there were too few trials per participant to compare EEG activity occurring during looks without any gaze shifts (*SI Appendix, Fig. S8*). A higher proportion of adult-led looks involved looks to other objects and the partner in the 2,000-ms time window, an effect that was driven by a greater number of object looks after the onset of adult-led attention (*SI Appendix, Fig. S8*). When the after look time period was broken down into 1,000-ms intervals, however, the difference between infant-led and adult-led looks was only seen in the first 1,000 ms after look onset (*SI Appendix, Figs. S9 and S10*), and a high proportion (>70%) of infant- and adult-led looks to mutual attention did not contain any object or partner looks (*SI Appendix, Fig. S8*). We nevertheless conducted a secondary analysis, excluding EEG activity for each look, in the time after infants shifted their attention away from the target object, toward another object/ the partner in the 2,000 ms after look onset. Two-dimensional cluster-based permutation analysis again revealed a significant positive cluster with an average frequency of 7 Hz (ranging 5 to 9 Hz), 104 to 1,994 ms after look onset ($P = 0.01$; *SI Appendix, Fig. S11*).

For comparison with the behavioral analysis presented in *Ostensive signals and infant attention*, *SI Appendix, Fig. S12* shows EEG activity over the same time windows (+5,000 ms; *SI Appendix, Fig. S12*). Two-dimensional cluster-based permutation analysis revealed no significant clusters of time*frequency points, comparing between infant-led looks and adult-led looks to mutual attention.

Followed vs. not followed infant-led attention.

Ostensive signals and infant attention. We examined whether ostensive signals differed between infant-led looks that were followed vs. not followed by their adult partner in the time after look onset. No significant difference in the likelihood of partner looks was observed, but there was a significant increase in the likelihood of infant vocalizations following infant-led nonmutual attention, 3 s after look onset (Fig. 4D). Again, this effect is likely driven by the significant reduction in infant vocalizations from baseline following infant-led looks to mutual

attention (Fig. 4D). Infant-led looks to nonmutual attention lasted a significantly shorter amount of time compared to infant-led looks to mutual attention, and this difference was marked [$t(36) = 6.84$, $P < 0.001$, $d = 1.13$; Fig. 4E]. Indeed, mutual attention extended infant attention irrespective of whether the attention episode was adult- or infant-led, with adult-led mutual attention episodes also lasting significantly longer compared to infant-led nonmutual attention [$t(36) = 8.25$, $P < 0.001$, $d = 1.36$; Fig. 4E].

Neural oscillatory activity. Corresponding to the significantly shorter object looks during episodes of nonmutual attention, infant-led looks to nonmutual attention included significantly fewer looks that lasted the whole 2,000 ms after look onset compared to infant-led looks to mutual attention, resulting in more looks containing object looks and looks to partner combined, both 0 to 1,000 ms and 1,000 to 2,000 ms after look onset (*SI Appendix, Figs. S8–S10*). Due to there being so few infant-led looks resulting in nonmutual attention that lasted the whole 2,000-ms time period (<50%; *SI Appendix, Fig. S8*), cluster-based permutation comparing infant-led looks to mutual and nonmutual attention was excluded from analysis.

Summary. Consistent with our hypothesis, infant-led mutual attention episodes were accompanied by significantly greater alpha desynchronization after look onset compared with adult-led mutual attention (*Neural oscillatory activity*). Against our predictions, infants also showed some decrease in their vocalizations after infant-led looks to mutual attention, compared to adult-led looks and infant-led looks to nonmutual attention, corresponding to a marked decrease from baseline after infant-led looks to mutual attention (*Ostensive signals and infant attention*). No differences in partner looks were observed (*Ostensive signals and infant attention*).

Discussion

This study investigated whether infants play a proactive role in creating episodes of joint attention during naturalistic tabletop play. In contrast to the results observed using structured, experimental paradigms (4, 9, 10), our results suggested that, in free-flowing interaction, 12-mo-old infants do not readily use their gaze or vocalize before an infant-initiated mutual attention episode; the occurrence of these behaviors throughout the interaction was generally low (Fig. 1 *A* and *B*). Although a significant difference in the probability of partner looks 1 s before look onset was identified, baseline comparisons indicated that this was driven by a reduction in infant looks to their partner before adult-initiated

looks rather than an increase before infant-initiated looks (Fig. 2A). Corresponding to the behavioral findings, and against what would be expected if infants were attempting to proactively drive their caregiver's attention toward an object, EEG activity at theta frequencies (3 to 6 Hz) did not increase in the 2 s before infant-led looks to mutual attention compared to adult-led looks: cluster-based permutation analysis revealed no significant clusters at any frequency band investigated (Fig. 3C).

Contrary to our prediction that infants' proactive engagement with their partner would affect whether a look was followed by the adult, no differences were identified between infant-led looks in ostensive signals (Fig. 2 C and D), or EEG activity (Fig. 3), before look onset. Taken together, the combination of our neural and behavioral results is inconsistent with the idea that infants routinely exert active and intentional control over the allocation of their attention where they lead their partner's attention and could suggest that similar processes drive infant attention when leading a mutual attention episode and joining the attentional focus of their partner (i.e., adult-led attention).

The null findings reported here are unlikely to be driven by eye movement-related artifact, introduced by temporally variable shifts in infant looking in the time before each look onset. Eye movement artifacts were removed using ICA decomposition, and although this does not remove all artifacts introduced to the EEG signal (44), we also show that each look type was equally affected by object and partner looks occurring in the 2,000 ms preceding look onset (*SI Appendix, Fig. S2*). The large sample size included here, particularly for infant EEG research (45), will have also increased signal-to-noise ratio in our data. It is also unlikely that this effect is driven by removal of neural activity during ICA decomposition: the algorithm used to reject ICA components during preprocessing (*Materials and Methods*) has been shown to be successful in retaining neural signal, especially in comparison to traditional manual rejection techniques (44). Furthermore, the secondary analysis conducted, excluding infant neural activity within each epoch, where the infant was not continuously focused on one object/the partner before look onset also showed no increase in theta power in the time before infant-led looks to mutual attention (*SI Appendix, Fig. S5*).

Although infant-led episodes of mutual attention did not appear proactively driven, infants were nevertheless sensitive to whether their look was followed by the adult. In line with hypotheses, in the time period after look onset, a significant decrease in EEG activity was observed over fronto-central electrodes in the alpha band (7 Hz), after infant-led looks to mutual attention, compared to adult-led looks (Fig. 5C). This finding reflects the pattern of neural activity observed in infants while watching the predicted outcome of another person's goal-directed behavior (39–41) and is consistent with previous experimental work showing reduced alpha activity where infant gaze was contingently responded to by a video-recorded experimenter (42). Thus, the reduction in EEG activity after infant-led looks to mutual attention may be interpreted as a neural marker of predictive processing during online social interaction, with infants predicting and encoding the behavioral contingency of their partner where they lead a look toward an object, and their partner follows. Against hypotheses, however, infants did not show an increase in looking to their partner in the time after look onset, suggesting that the anticipated contingency of their caregiver was not realized by the infant through observing partner behaviors signaling intention to share attention (7).

A possible interpretation of our findings is that, rather than shared intentionality, inter-dyadic coordination is largely achieved and perceived by the infant through attending toward their

partner's sensorimotor behaviors. In line with previous findings in naturalistic studies, infants did not readily follow their partner's gaze (15). In fact, infants looked to objects more (rather than to their partner) in the 1-s time period before adult-initiated looks. This is consistent with Yu and Smith's observation that moments infants join their partner's attention are driven by the partner's manual activity on objects (15, 19). The neural analyses of the current study, that show no increase in endogenous oscillatory activity before infant-led looks, relative to adult-led looks, suggest that similar, external inputs might also drive infant attention where they lead a look toward an object. As well as overt behaviors such as object manipulations and gestural communication, other sensory inputs could also influence shifts in infant gaze. For example, in very early face-to-face interactions, salient events such as pauses in adult vocalizations and changes in the fundamental frequency of their voice modulate infant attention toward and away from the partners' face (46, 47). In the current study, analysis of caregivers' ostensive signaling revealed that partner looks increased and vocalizations decreased in the time before infant-led looks to mutual attention, compared to adult-led looks, and infant-led looks to nonmutual attention (*SI Appendix, Fig. S13*).

Entrainment to the low-level sensorimotor dynamics of shared interactions (3) could be the mechanism through which infants *perceive* the behavioral contingency of their communicative partner, suggested by the alpha suppression observed after infant-led looks to mutual attention. Research into action-oriented predictive processing suggests that motor intentions actively elicit active predictions about the ongoing consequences of our own actions (48, 49). Perhaps similar processes operate across the dyad during early behavioral coordination, with the infant anticipating the effect of their own action on the behavior of their partner (27, 28). Again, as well as overt manual behaviors (15), other fast-changing cues such as temporal and spectral modulations in the partner's vocalizations could also signal behavioral contingency to the infant (46, 47). Interestingly, the probability of infants vocalizing falls below baseline in the time after infant-led looks to mutual attention. This finding is possibly indicative of the sensorimotor turn-taking processes occurring after infant-led looks to mutual attention, i.e., that infants are anticipating the behavioral response of their partner in the time after they shift their attention toward a new object.

That said, while the findings of the current analysis suggest that infants do not routinely show signs of proactively leading their partner's attention during shared interaction, it is still possible that moments of proactive engagement by the infant are infrequent but nevertheless important to the ongoing interactive exchange (50). The results of the behavioral analysis, investigating ostensive signals occurring in the time before and after infant-led looks to mutual and nonmutual attention, show that the probability of these behaviors occurring before infants lead a look to their partner is low but not absent (Fig. 2 A–D and 4 A–D). An interesting question is whether moments ostensive signals do occur before infants initiate an attention episode with their partner are largely incidental, or whether all or some of these moments occur as a result of the infant attempting to actively engage their adult partner's attention but remain too rare to make a difference at the statistical level. The mechanisms through which infants perceive the behavioral contingency of their partner during active attention-sharing episodes, if they occur, could be functionally different to those engaged during externally driven attention.

Our findings have important implications for how we view and understand the learning processes involved in early joint attention development. Associative learning accounts postulate that infants learn about their environment, and how to act on it, through

repeated reinforcement, where the value given to an action is based on previous experience of how that action affected the environment (51–53). In the context of social interaction, infant behavior is assigned meaning by the adult through consistent and contingent behavioral feedback. Over time, these statistical regularities form the basis for infant representations about the intentions of others and how their own intentionally motivated behaviors affect those of their partner (53). In line with this, the current EEG findings suggest that infants predict and encode the behavioral contingency of their partner to their own actions before they show signs of intentionally initiating joint attention episodes in a routine manner. Caregiver responses to infrequent moments that infants engage in proactive attention sharing may therefore be particularly important to the development of infants' representations about their own intentionally motivated behavior and, over time, increase the extent to which infants use these behaviorally reinforced cues to proactively direct the attention of their partner (53).

This perspective has potential implications for current theories of how infants begin to acquire a language system. A popular view has been that, it is only once infants are able to establish and understand a joint attentional frame, or "common ground," between themselves, an object, and their partner, they can begin to engage with and learn from the pragmatic and referential aspects of shared communication (1, 4, 54). Our findings, however, consistent with an associative learning framework of joint attention development (53), suggest that before infants routinely engage in triadic forms of shared attention, infant attunement to their partner's sensorimotor behaviors, and the timing of adult inputs as a function of infant behaviors may already be contributing toward infants developing understanding of the principles underlying communication. Recent behavioral work has, in fact, suggested that infants' engagement with objects at the time object labels are presented by their communicative partner could go some way in solving the problem of referential ambiguity (18, 55). Combining neural and behavioral methods to explore how infant attunement to action-generated contingencies during naturalistic free-flowing interactions supports early language acquisition should be a key focus for future research.

While this study shows how infant neural activity changes around moments of infant- vs. adult-led episodes of mutual attention during naturalistic interactions, predictive encoding models investigating the dynamic relationship between infant attention, interdyadic behavior, and infant neural activity should be a next step (56). Neural tracking of auditory information to controlled experimental stimuli has been shown in both adults (57) and more recently infants (58, 59). Whether and how infants' neural activity dynamically responds to modulations in their partners' behaviors, including features and the timing of caregiver vocalizations, manual activity, and bodily movement, and how this associates with the timing of infant- and adult-led episodes of joint action are yet to be investigated. Examining these questions developmentally will be integral to understanding the development of intentional communication in infancy and in identifying atypical trajectories (60).

Our use of naturalistic data is a limitation, as well as a strength, as we were unable to control for how much infants moved their attention between objects in the time before and after look onsets. This not only introduces artifact to the EEG signal but also means that the extent to which oscillatory activity is influenced by object processing differs between looks. However, we showed that the number of object and partner looks did not differ in the 2,000 ms before look onset for either comparison, and differed only in the first 1,000 ms after look onset, comparing infant- and adult-led

looks to mutual attention (*SI Appendix*, Figs. S2, S8, and S9). Our secondary analyses, excluding infant EEG activity where the infant was not continually focused on one object or the partner before/after look onset, also showed similar differences in patterns of neural oscillatory activity between looks reported in the main text (*SI Appendix*, Figs. S5 and S11). Increased gaze shifts after infant-led looks to nonmutual attention did, however, mean that we were unable to compare infant-initiated looks to mutual and nonmutual attention in the time after look onset. Employing continuous methods of analysis to naturalistic data would overcome this issue (56).

The ability to engage in reciprocally mediated joint attention toward the end of the first year is catalytic to developments in language and social cognition (4, 18). The findings reported here suggest that at 10 to 12 mo, infants are not yet predominantly proactive in creating and maintaining episodes of joint attention with their adult partner. They are, however, sensitive to whether their behavior is contingently responded to, potentially forming the basis for the emergence of intentionally mediated communication.

Materials and Methods

Participants. Fifty-eight caregiver–infant dyads took part in the study; 37 participants contributed usable data [13 excluded due to recording error, 2 excluded due to infant fussiness, and 6 excluded due to poor quality infant EEG (see the *Artifact Rejection and Preprocessing* section for more information on EEG exclusion criteria)]. The final sample included 18 females and 19 males; mean age, 11.12 mo ($SD = 1.33$). All caregivers were females. Participants were recruited through baby groups and children's centers in the boroughs of Newham and Tower Hamlets, as well as through online platforms such as Facebook, Twitter, and Instagram. Written informed consent was obtained from all participants before taking part in the study, and consent to publish was obtained for all identifiable images used. All experimental procedures were reviewed and approved by the University of East London Ethics Committee.

Experimental Setup. Caregivers and infants were seated facing each other on opposite sides of a 65-cm wide table. Infants were seated in a high chair, within easy reach of the toys (Fig. 6C). The shared toy play comprised two sections, with a different set of toys in each section, each lasting ~5 min each. Two different sets of three small, age-appropriate toys were used in each section; this number was chosen to encourage caregiver and infant attention to move between the objects while leaving the table uncluttered enough for caregiver and infant gaze behavior to be accurately recorded (cf. ref. 19).

At the beginning of the play session, a researcher placed the toys on the table, in the same order for each participant, and asked the caregiver to play with their infant just as they would at home. Two researchers stayed behind a screen out of view of caregiver and infant, except for the short break between play sessions. The mean length of joint toy play recorded combining the first and second play sections was 9.92 min ($SD = 2.31$).

Equipment. EEG signals were recorded using a 32-channel BioSemi gel-based ActiveTwo system with a sampling rate of 512 Hz with no online filtering using ActiView software. The interaction was filmed using three Canon LEGRIA HF R806 camcorders recording at 50 fps. Caregiver and infant vocalizations were also recorded throughout the play session using a ZOOM H4n Pro Handy Recorder and Sennheiser EW 112P G4-R receiver.

Two cameras faced the infant: one placed on the left of the caregiver and one on the right (Fig. 6C). Cameras were placed so that the infant's gaze and the three objects placed on the table were clearly visible, as well as a side view of the caregiver's torso and head. One camera faced the caregiver, positioned just behind the left or right side of the infant's high chair (counterbalanced across participants). One microphone was attached to the caregiver's clothing and the other to the infant's high chair.

Caregiver and infant cameras were synchronized to the EEG via radio frequency (RF) receiver LED boxes attached to each camera. The RF boxes simultaneously received trigger signals from a single source (computer running MATLAB) at the beginning of each play section, and concurrently emitted light impulses, visible in

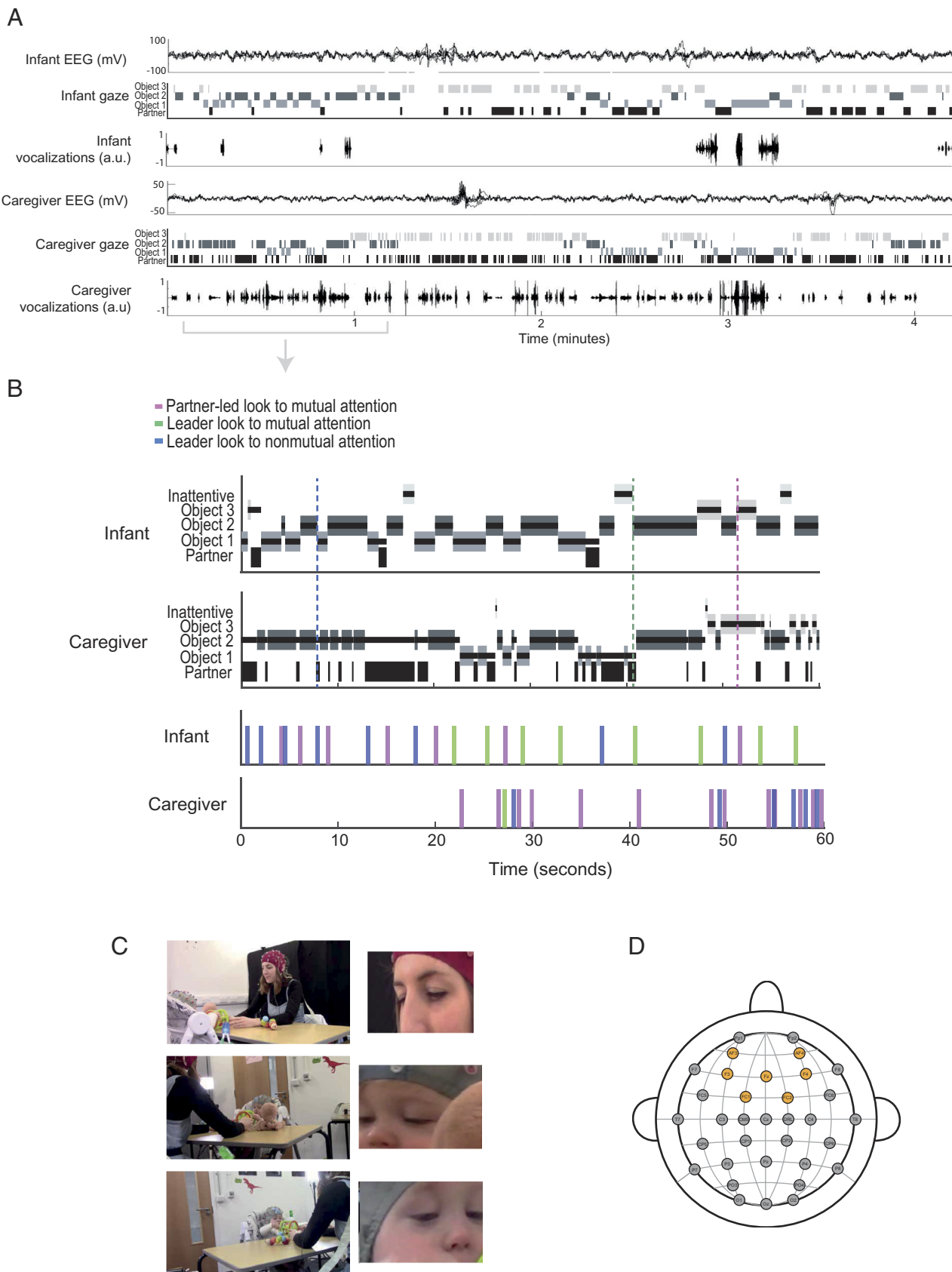


Fig. 6. Example data collected during one 5-min interaction for one dyad and camera angles used for coding and EEG montage. (A) Raw data sample showing (from Top) infant EEG over fronto-central electrodes, after preprocessing, infant gaze behavior, infant vocalizations, adult EEG over fronto-central electrodes, adult gaze behavior, and adult vocalizations. (B) Example of interpolated looks (thin black lines) superimposed on caregiver and infant looking behavior before interpolation (thick gray lines). Colored dashed lines indicate examples of different look types in the infant gaze time series (Top). Spike trains for infant and caregiver looks colored according to look type (Bottom). (C) Example camera angles for caregiver and infant (Right and Left), as well as zoomed-in images of caregiver and infant faces, used for coding. (D) Topographical map showing electrode locations on the BioSemi 32-cap fronto-central electrodes included in the infant time–frequency analysis are highlighted in orange (AF3, AF4, FC1, FC2, F3, F4, and Fz).

each camera. Microphone data were synchronized with the infants' video stream via a xylophone ding recorded in the infant camera and both microphones, which was hand-identified in the recordings by trained coders. All systems were extensively tested and found to be free of latency and drift between EEG, camera, and microphone to an accuracy of ± 20 ms.

Video Coding. The visual attention of caregiver and infant was manually coded using custom-built MATLAB scripts that provided a zoomed-in image of caregiver and infant faces (see Fig. 6C). Coders indicated the start frame (i.e., to the closest 20 ms at 50 fps) that caregiver or infant looked to one of the three objects, to their partner, or looked away from the objects or their partner (i.e. became inattentive). Partner looks included all looks to the partner's face; looks to any other parts of the body or the cap were coded as inattentive. Periods where the researcher was within camera frame were marked as uncodable, as well as instances where the caregiver or infant gaze was blocked or obscured by an object, or their eyes were outside the camera frame. Video coding was completed by two coders, who were trained by the first author. Interrater reliability analysis on 10% of coded interactions (conducted on either play section 1 or play section 2), dividing data into 20-ms bins, indicated strong reliability between coders ($\kappa = 0.9$ for caregiver coding and $\kappa = 0.8$ for infant coding).

Vocalization Coding. The onset and offset times of caregiver and infant vocalizations were coded using custom-built MATLAB scripts that allowed coders to identify the onset and offset of a vocalization based on the spectrogram, as well as auditory sound. A vocalization was defined as a continuous sound produced by the caregiver or infant, with a pause less than 500 ms. Due to the labor-intensive nature of the vocalization coding, vocal coding was completed for a subsample of the caregiver-infant dyads ($n = 19$). Interrater reliability on 10% of coded interactions (conducted on either play section 1 or play section 2), dividing data into 1-ms bins, again indicated strong reliability between coders ($\kappa = 0.8$).

Behavioral Look Extraction and Analysis.

Data preprocessing. The aim of our analysis was to identify moments where the infant's attention transitioned from one play object to another and to examine whether the infant or the caregiver initiated the transition. Before doing this, however, we first interpolated through infant and caregiver looks to their partner. This is because, as shown in Fig. 6 A and B, during periods of concurrent looking toward an object, caregivers, and, to a lesser extent, infants, alternated their attention frequently between the object and their partner. Without interpolation, each subsequent look back to the object would be classified as a separate follower look to the object. This procedure thus allowed us to accurately identify moments in the interaction where the infant was leading and following their partner's attention while considering the dynamic nature of joint attention documented in previous studies (19).

Interpolation involved identifying moments where the caregiver or infant looked up to their partner and then interpolating through that look, so that the partner look became an extension of the preceding object look. No threshold was set for interpolation: a new look was considered to have started at the beginning of each new object look (Fig. 6B). After interpolation, the first and last frames of all attention episodes were extracted. Infant object looks were categorized into adult-led and infant-led looks. Infant-led looks were subdivided into two further categories: infant-led looks to mutual and nonmutual attention (Table 1 for description of each look category). Looks that followed or preceded uncodable gaze behavior were excluded from analysis, as well as leader looks where the partner's gaze in the time after look onset preceded an uncodable period.

Cluster-based permutation analysis—behavioral data. To test for significant differences in the likelihood of ostensive signals during the time periods before and after infant-led and adult-led looks, a permutation-based temporal clustering analysis was conducted (61). This approach controls for familywise error rate using a nonparametric Monte Carlo method. A full description of the cluster-based permutation analysis is given in [SI Appendix, Supplementary Methods](#).

Infant EEG Analysis.

Artifact rejection and preprocessing. A fully automatic artifact rejection procedure including ICA was adopted, following procedures from commonly used toolboxes for EEG preprocessing in adults (62, 63) and infants (64, 65) and

Table 1. Definition of infant attention episode categories

Attention episode	Definition
Adult-led looks to mutual attention	The start of the attention episode was taken from the frame that the infant first shifted their gaze toward an object that the adult was already looking toward at any point in the time that the adult was still attending toward the object.
Infant-led looks to mutual attention	The start of the attention episode was taken from the frame that the infant first shifted their gaze toward an object that the adult was not already looking at, and the adult subsequently joined the infant's gaze toward the object at any point in the time that the infant was still attending toward the object.
Infant-led looks to nonmutual attention	The start of the attention episode was taken from the frame that the infant first shifted their gaze toward an object that the adult was not already looking at, and the adult did not follow the infant's look toward the object at any point in the time that the infant was still attending toward the object.

optimized and tested for use with our naturalistic infant EEG data (44, 66, 67). Full details of EEG artifact rejection and preprocessing are given in [SI Appendix, Supplementary Methods](#).

Time-frequency analysis. Each infant look onset was identified in the EEG signal, and activity occurring 2,500 ms before to 2,500 ms after look onset extracted across all channels. An additional 200 ms was also extracted immediately prior to this segment to serve as the prelook baseline. Only look epochs with 25% or fewer data points excluded during artifact rejection were included in analysis, and missing data points were set to NaN.

Time-frequency decomposition was conducted on each look epoch via continuous Morlet wavelet convolution, whereby the EEG signal at each channel was convolved with Gaussian-windowed complex sine waves, ranging from 1 to 16 Hz, in linearly spaced intervals. This frequency range was selected as the frequency range least sensitive to movement artifacts inherent in naturalistic infant EEG, which affects both low- (<2 Hz) and high-frequency (>16 Hz) activity (66). The width of the Gaussian was set to seven cycles. Before wavelet convolution, the epoched data were reshaped into continuous data and afterward transformed back to individual epochs. To remove distortion introduced by wavelet convolution, the first and last 500 ms of each epoch were chopped off, so that the epochs were 4,200 ms in length. After convolution, power was extracted as the absolute value squared, resulting from the complex signal, before averaging power values at each time point over all looks. The condition-specific baseline period used was 2,200 to 2,000 ms before look onset. Averaged power time series occurring 2,000 ms before and after look onset were normalized by transforming the baseline-corrected signal to a decibel (dB) scale (68).

Cluster-based permutation analysis—EEG data. Two approaches were used for analyzing the EEG data. First, two-dimensional (frequency x time) clusters were calculated based on data collapsed in topographical space over fronto-central electrodes (Fig. 6D). Second, three-dimensional (frequency x time x electrode) clusters were calculated based on the entire data. For the first analysis, normalized power was averaged over fronto-central electrodes (AF3, AF4, FC1, FC2, F3, F4, and Fz; see Fig. 6D for locations) and compared between looks. This electrode cluster was chosen based on previous infant literature (69). Only participants contributing five usable trials or more to both look categories, in one comparison, were included in each analysis (see [SI Appendix, Fig. S1](#), for details on the number of epochs included before and after artifact rejection for each type of attention episode). Second, in order to examine how the distribution of results varied topographically over the brain, an additional three-dimensional cluster-based permutation

analysis was conducted to examine time–frequency–electrode space for clusters of significant data points. See [SI Appendix, Supplementary Methods](#) for a full description of the two-dimensional and three-dimensional cluster-based permutation analyses.

Data, Materials, and Software Availability. De-identified EEG and behavioural time-series data analysed for this paper have been deposited in the OSF (<https://osf.io/35sry/>) (70). Due to the personally identifiable nature of the video recordings (interactions between caregivers and their infants), the raw video data

and microphone recordings are not publicly accessible. For any queries, please contact the first author: u1920558@uel.ac.uk.

ACKNOWLEDGMENTS. This research was funded by The Leverhulme Trust. Thank you to Lynne Murray for reading through and giving detailed comments on earlier versions of the manuscript. Thanks to Emily Greenwood and Dean Matthews for help with data coding. Thanks to members of the UEL BabyDev Lab for comments and discussions on earlier drafts of this manuscript and to all participating children and caregivers.

1. M. Tomasello, M. Carpenter, J. Call, T. Behne, H. Moll, Understanding and sharing intentions: The origins of cultural cognition. *Behav. Brain Sci.* **28**, 675–691 (2005).
2. C. D. Frith, U. Frith, Social cognition in humans. *Curr. Biol.* **17**, R724–R732 (2007).
3. N. Sebanz, G. Knoblich, Prediction in joint action: What, when, and where. *Topics in Cognitive Sci.* **1**, 353–367 (2009).
4. M. Carpenter, K. Nagell, M. Tomasello, G. Butterworth, C. Moore, Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Dev.* **63**, 1–143 (1998).
5. E. Donnellan, C. Bannard, M. L. McGillion, K. E. Slocombe, D. Matthews, Infants' intentionally communicative vocalizations elicit responses from caregivers and are the best predictors of the transition to language: A longitudinal investigation of infants' vocalizations, gestures and word production. *Dev. Sci.* **23**, e12843 (2020).
6. J. M. Iverson, S. Goldin-Meadow, Gesture paves the way for language development. *Psychol. Sci.* **16**, 367–371 (2005).
7. B. Sipoova, M. Carpenter, A new look at joint attention and common knowledge. *Cognition* **189**, 260–274 (2019).
8. M. Tomasello, M. Carpenter, Shared intentionality. *Dev. Sci.* **10**, 121–125 (2007).
9. U. Liszkowski, M. Carpenter, A. Henning, T. Striano, M. Tomasello, Twelve-month-olds point to share attention and interest. *Dev. Sci.* **7**, 297–307 (2004).
10. U. Liszkowski, K. Albrecht, M. Carpenter, M. Tomasello, Infants' visual and auditory communication when a partner is or is not visually attending. *Cognition* **108**, 732–739 (2008).
11. K. Begus, V. Southgate, "Curious learners: How infants' motivation to learn shapes and is shaped by infants' interactions with the social world" in *Active Learning from Infancy to Childhood*, M. M. Saylor, P. A. Ganea, Eds. (Springer International Publishing, 2018), pp. 13–37.
12. Á. M. Kovács, T. Tazuin, E. Téglás, G. Gergely, G. Csibra, Pointing as epistemic request: 12-month-olds point to receive new information. *Infancy* **19**, 543–557 (2014).
13. K. Begus, V. Southgate, Infant pointing serves an interrogative function. *Dev. Sci.* **15**, 611–617 (2012).
14. L. Goupil, M. Romand-Monnier, S. Kouider, Infants ask for help when they know they don't know. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 3492–3496 (2016).
15. C. Yu, L. B. Smith, Joint attention without gaze following: human infants and their parents coordinate visual attention to objects through eye-hand coordination. *PLoS One* **8**, e79659 (2013).
16. S. V. Wass, M. Whitehorn, I. Marriott Haresign, E. Phillips, V. Leong, Interpersonal neural entrainment during early social interaction. *Trends in Cognitive Sci.* **24**, 329–342 (2020).
17. S. V. Wass *et al.*, Infants' visual sustained attention is higher during joint play than solo play: Is this due to increased endogenous attention control or exogenous stimulus capture? *Dev. Sci.* **21**, e12667 (2018).
18. C. Yu, L. B. Smith, Embodied attention and word learning by toddlers. *Cognition* **125**, 244–262 (2012).
19. C. Yu, L. B. Smith, Multiple sensory-motor pathways lead to coordinated visual attention. *Cognitive Sci.* **41**, 5–31 (2017).
20. S. A. Custode, C. Tamis-LeMonda, Cracking the code: Social and contextual cues to language input in the home environment. *Infancy* **25**, 809–826 (2020).
21. J. M. Franchak, K. S. Kretch, K. C. Soska, K. E. Adolph, Head-mounted eye tracking: A new method to describe infant looking: head-mounted eye tracking. *Child Dev.* **82**, 1738–1750 (2011).
22. K. Begus, E. Bonawitz, The rhythm of learning: Theta oscillations as an index of active learning in infancy. *Dev. Cognitive Neurosci.* **45**, 100810 (2020).
23. E. Orekhova, T. A. Stroganova, I. N. Posikera, Theta synchronization during sustained anticipatory attention in infants over the second half of the first year of life. *Int. J. Psychophysiol.* **32**, 151–172 (1999).
24. W. Xie, B. M. Mallin, J. E. Richards, Development of infant sustained attention and its relation to EEG oscillations: An EEG and cortical source analysis study. *Dev. Sci.* **21**, e12562 (2018).
25. E. Orekhova, T. Stroganova, I. Posikera, M. Elam, EEG theta rhythm in infants and preschool children. *Clin. Neurophysiol.* **117**, 1047–1062 (2006).
26. S. V. Wass *et al.*, Parental neural responsiveness to infants' visual attention: How mature brains influence immature brains during social interaction. *PLoS Biol.* **16**, e2006328 (2018).
27. A. F. de C. Hamilton, Hyperscanning: Beyond the hype. *Neuron* **109**, 404–407 (2021).
28. I. Konvalinka *et al.*, Frontal alpha oscillations distinguish leaders from followers: Multivariate decoding of mutually interacting brains. *NeuroImage* **94**, 79–88 (2014).
29. U. Hasson, C. D. Frith, Mirroring and beyond: Coupled dynamics as a generalized framework for modelling social interactions. *Phil. Trans. R. Soc. B* **371**, 20150366 (2016).
30. G. M. Mason, F. Kirkpatrick, J. A. Schwade, M. H. Goldstein, The role of dyadic coordination in organizing visual attention in 5-month-old infants. *Infancy* **24**, 162–186 (2019).
31. G. M. Mason, *Investigating Dyadic Social Coordination and Infant Attention in Typical and Atypical Development* (Cornell University, 2018).
32. J. L. Miller, J. Gros-Louis, Socially guided attention influences infants' communicative behavior. *Infant Behav. Dev.* **36**, 627–634 (2013).
33. J. L. Miller, E. Hurdish, J. Gros-Louis, Different patterns of sensitivity differentially affect infant attention span. *Infant Behav. Dev.* **53**, 1–4 (2018).
34. W. Klimesch, P. Sauseng, S. Hanslmayr, EEG alpha oscillations: The inhibition-timing hypothesis. *Brain Res. Rev.* **53**, 63–88 (2007).
35. G. Thut, Band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *J. Neurosci.* **26**, 9494–9502 (2006).
36. J. M. Kilner, C. Vargas, S. Duval, S.-J. Blakemore, A. Sirigu, Motor activation prior to observation of a predicted movement. *Nat. Neurosci.* **7**, 1299–1301 (2004).
37. R. Hari *et al.*, Activation of human primary motor cortex during action observation: A neuromagnetic study. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 15061–15065 (1998).
38. S. D. Muthukumaraswamy, B. W. Johnson, N. A. McNair, Mu rhythm modulation during observation of an object-directed grasp. *Cognitive Brain Res.* **19**, 195–201 (2004).
39. C. D. Monroy, M. Meyer, L. Schröder, S. A. Gerson, S. Hunnius, The infant motor system predicts actions based on visual statistical learning. *NeuroImage* **185**, 947–954 (2019).
40. M. Meyer, R. Braukmann, J. C. Stapel, H. Bekkering, S. Hunnius, Monitoring others' errors: The role of the motor system in early childhood and adulthood. *Br. J. Dev. Psychol.* **34**, 66–85 (2016).
41. V. Southgate, M. H. Johnson, I. E. Karoui, G. Csibra, Motor system activation reveals infants' on-line prediction of others' goals. *Psychol. Sci.* **21**, 355–359 (2010).
42. H. Rayson, J. J. Bonaiuto, P. F. Ferrari, B. Chakrabarti, L. Murray, Building blocks of joint attention: Early sensitivity to having one's own gaze followed. *Dev. Cognitive Neurosci.* **100631** (2019).
43. S. Hoehl, C. Michel, V. M. Reid, E. Parise, T. Striano, Eye contact during live social interaction modulates infants' oscillatory brain activity. *Soc. Neurosci.* **9**, 300–308 (2014).
44. I. Marriott Haresign *et al.*, Automatic classification of ICA components from infant EEG using MARA. *Dev. Cognitive Neurosci.* **52**, 101024 (2021).
45. V. Noreika, S. Georgieva, S. Wass, V. Leong, 14 challenges and their solutions for conducting social neuroscience and longitudinal EEG research with infants. *Infant Behav. Dev.* **58**, 101393 (2020).
46. C. L. Crown, S. Feldstein, M. D. Jasnow, B. Beebe, J. Jaffe, The cross-modal coordination of interpersonal timing: six-week-olds infants' gaze with adults' vocal behavior. *J. Psycholinguistic Res.* **31**, 1–23 (2002).
47. D. N. Stern, S. Spieker, K. MacKain, Intonation contours as signals in maternal speech to prelinguistic infants. *Dev. Psychol.* **18**, 727 (1982).
48. K. J. Friston, Waves of prediction. *PLoS Biol.* **17**, e3000426 (2019).
49. A. Clark, Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* **36**, 181–204 (2013).
50. L. Murray *et al.*, The functional architecture of mother-infant communication, and the development of infant social expressiveness in the first two months. *Sci. Rep.* **6**, 1–9 (2016).
51. P.-Y. Oudeyer, L. B. Smith, How evolution may work through curiosity-driven developmental process. *Top. Cogn. Sci.* **8**, 492–502 (2016).
52. G. O. Deák, A. M. Krasno, J. Triesch, J. Lewis, L. Sepeta, Watch the hands: Infants can learn to follow gaze by seeing adults manipulate objects. *Dev. Sci.* **17**, 270–281 (2014).
53. L. B. Smith, C. Breazeal, The dynamic lift of developmental process. *Dev. Sci.* **10**, 61–68 (2007).
54. E. Lieven, Usage-based approaches to language development: Where do we go from here? *Lang. Cogn.* **8**, 346–368 (2016).
55. C. Yu, Y. Zhang, L. K. Slone, L. B. Smith, The infant's view redefines the problem of referential uncertainty in early word learning. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2107019118 (2021).
56. S. Jessen, J. Obleser, S. Tune, Neural tracking in infants – an analytical tool for multisensory social processing in development. *Dev. Cognit. Neurosci.* **52**, 101234 (2021).
57. E. M. Zion Golumbic *et al.*, Mechanisms underlying selective neuronal tracking of attended speech at a "Cocktail Party". *Neuron* **77**, 980–991 (2013).
58. A. Attaheri, Delta- and theta-band cortical tracking and phase-amplitude coupling to sung speech by infants. *NeuroImage*, in press.
59. M. Kalashnikova, V. Peter, G. M. Di Liberto, E. C. Lalor, D. Burnham, Infant-directed speech facilitates seven-month-old infants' cortical tracking of speech. *Sci. Rep.* **8**, 13745 (2018).
60. E. A. Gaffan, C. Martins, S. Healy, L. Murray, Early social experience and individual differences in infants' joint attention. *Soc. Dev.* **19**, 369–393 (2010).
61. E. Maris, R. Oostenveld, Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* **164**, 177–190 (2007).
62. T. Mullen, "CleanLine EEGLAB plugin" in *San Diego, CA, Neuroimaging Informatics Tools and Resources Clearinghouse (NITRC)* (2012).
63. N. Bigdelly-Shamlo, T. Mullen, C. Kothe, K.-M. Su, K. A. Robbins, The PREP pipeline: standardized preprocessing for large-scale EEG analysis. *Front. Neuroinform.* **9**, 16 (2015).
64. L. J. Gabard-Durnam, A. S. Mendez Leal, C. L. Wilkinson, A. R. Levin, The Harvard automated processing pipeline for electroencephalography (HAPPE): Standardized processing software for developmental and high-artifact data. *Front. Neurosci.* **12**, 97 (2018).
65. R. Debnath *et al.*, The Maryland analysis of developmental EEG (MADE) pipeline. *Psychophysiology* **57**, e13580 (2020).
66. S. Georgieva *et al.*, Toward the understanding of topographical and spectral signatures of infant movement artifacts in naturalistic EEG. *Front. Neurosci.* **14**, 352 (2020).
67. I. Marriott Haresign, E. Phillips, M. Whitehorn, L. Goupil, S. V. Wass, Using dual EEG to analyse event-locked changes in child-adult neural connectivity (2021). *bioRxiv* [Preprint]. <https://doi.org/10.1101/2021.06.15.448573>. Accessed 12 December 2021.
68. M. X. Cohen, *Analyzing Neural Time Series Data: Theory and Practice* (The MIT Press, 2014).
69. E. K. Braithwaite, E. J. H. Jones, M. H. Johnson, K. Holmboe, Dynamic modulation of frontal theta power predicts cognitive ability in infancy. *Dev. Cognit. Neurosci.* **45**, 100818 (2020).
70. E. Phillips, Proactive or reactive? Neural oscillatory insight into the leader-follower dynamics of early infant-caregiver interaction. *Open Science Framework*. <https://osf.io/35sry/>. Deposited 20 March 2023.