SHORT COMMUNICATION
Neural mechanisms of joint attention in infancy

Tricia Striano,[1,2,3] Vincent M. Reid[2] and Stefanie Hoehl[2]
[1]Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstrasse 1a, 04103 Leipzig, Germany
[2]Neurocognition and Development Group, Center for Advanced Studies, University of Leipzig, Otto-Schill Strasse, Leipzig, Germany
[3]Department of Pediatrics and Kennedy Center for Human Development, Vanderbilt University, Nashville, TN, USA

Keywords: event-related potentials, infants, joint attention, live ERP paradigm, social cognition

Abstract

One of the key transitions in early cognitive development is from participating in face-to-face interactions to engaging in joint attention exchanges. It is known that the ability to jointly attend with another person to an object is essential for the development of abilities such as language in later life. Strikingly, little is known about the function of joint attention in infants in the first year. We developed a novel interactive-live paradigm to assess the neural mechanisms of joint attention in 9-month-old infants. An adult interacted with each infant, and infants' electrical brain activity was measured in two contexts. In the joint attention context, a live adult gazed at the infants' face and then to a computer displayed novel object. In the non-joint attention context the adult gazed only to the novel object. We found that the negative component of the infant event-related potential (ERP), a neural correlate indexing attentional processes, was enhanced in amplitude during the processing of objects when infants were engaged in a joint attention interaction compared to a non-joint attention interaction. These results suggest that infants benefit from joint attention interactions by focusing their limited attentional resources to specific aspects of the surrounding environment.

Introduction

From the first seconds of postnatal experience, infants prefer to look at human faces rather any other form of visual stimuli (Johnson & Morton, 1991; Farroni *et al*., 2002). This capacity to seek out and engage with others develops and changes remarkably during the first postnatal year. By the middle of the first year, infants begin to coordinate objects into their interactions with people (Striano & Bertin, 2005). One of the key transitions in early cognitive development is from having the capacity to participate in dyadic face-to-face interactions to systematically engage in joint attention exchanges. Joint attention is a triadic social interaction that involves monitoring: (i) another person's attention in relation to the self; (ii) an external object or event, and (iii) the other person's attention toward the same external object. The ability to jointly attend with another person has been shown to be crucial for the learning of language and for imitative learning (Baldwin, 1993). Given its essential function in the development of human social cognitive skills, joint attention has attracted much research in recent years.

Recently it has been demonstrated that by 3 months of age, infants are already able to discriminate between dyadic and joint attention interactions (Striano & Stahl, 2005). However, the function of infants' sensitivity to joint attention in terms of its impact on perceptual processing and cognitive abilities remains open.

In parallel and in isolation to research investigating joint attention, much research has been conducted into properties of infant electrophysiological responses to cognitive tasks. Most often these responses have been measured with event-related potentials (ERP), which is that electrical brain activity that is time locked to the onset of a stimulus (Rugg & Coles, 1995). In the past ten years, knowledge of how the functional brain develops has increased dramatically (Johnson, 2005). One component of the infant ERP that is well mapped in terms of cognitive properties is the mid-latency negative component, or Nc. The Nc occurs approximately 300–700 milliseconds after stimulus onset, is most prominent at fronto-central electrodes, and is thought to relate to the development of memory processes during the first 12 postnatal months (Webb *et al*., 2005). The component is thought to reflect attentional orienting to salient stimuli (Courchesne *et al*., 1981; Nelson, 1994) and/or a general attentional arousal (Richards, 2003a) as it is larger to infrequent than frequent stimuli (e.g, Courchesne *et al*., 1981), and is larger during periods of sustained attention as defined by heartrate (Richards, 2003b). Recently it has been suggested that the Nc can be localized to the anterior cingulate and other regions of the frontal cortex (Reynolds & Richards, 2005).

We tested the influence of joint attention interactions on 9-month-old infants' object processing. We predicted an enhanced Nc for objects that infants viewed in the context of joint attention compared to non-joint attention contexts.

Materials and methods

Participants

All research was conducted under institutional protocols, which were approved by the Ethical Committee of the Humboldt University of Berlin, with parental informed consent for each subject. Fifteen infants (seven males and eight females) with an average age of 9 months ± 12 days were tested. All infants were born full term (37–41 weeks)

Correspondence: Dr Tricia Striano, [1]Max Planck Institute for Human Cognitive and Brain Sciences, as above.
E-mail: striano@cbs.mpg.de

Received 23 December 2005, revised 13 March 2006, accepted 15 March 2006

and were in the normal range for birthweight. Another 52 infants were tested but were excluded from the final sample as a result of fussiness of the infant ($n = 10$), failing to reach the minimum requirements for adequate averaging of the ERP data ($n = 39$), or experimental error ($n = 3$). This relatively high dropout rate was due to task demands, and the inclusion of a trial only if an infant produced an overt gaze to the experimenters face and then to the screen. The minimum criteria for inclusion was ten trials per condition, however, each infant contributed 20–31 (mean of 23.3) trials to their average from a mean of 73.3 viewed presentations of the novel objects.

Procedure

An adult interacted with each infant, and infants' electrical brain activity was measured in two contexts. In the joint attention context, the adult gazed at the infant's face and then to a novel object that was displayed on a computer screen for one second (see Fig. 1). In the non-joint attention context the adult gazed only to the novel object, with the initial position of the adult's gaze fixed towards the monitor displaying the objects (see Fig. 2). The amount of adult vocalizations and positive affect was the same in both contexts. Within a period of 5 s, and for each trial included in analysis, infants were required to look at the adult's face and then to the novel object on the screen. Our criteria for further analysis ensured that only trials that elicited overt head movements were included, irrespective of condition. This effectively controlled the amount of motion for infant head turn.

Infants sat on their mother's lap in a dimly lit sound-attenuated and electrically shielded cabin, at a viewing distance of 90 cm away from a 70 Hz 17-inch stimulus monitor and 80 cm away from the experimenter's face. The experimenter's face was 40 cm away from the centre of the stimulus monitor. The experiment consisted of one block with 200 trials (100 joint attention, 100 no joint attention).

The two conditions were presented to the infant in a random order. However, the same condition was not presented three times consecutively and the number of presentations of each condition was balanced within every 20 presentations. Each object on the screen was presented for 1 s and subtended a visual angle of 4.3°. Each object was preceded by a small fixation object (visual angle of 3.8°), presented in the middle of the screen for 500 ms. This object was included in order to avoid infant muscle artefact related to head turning and eye movement artefacts in the data. For inclusion in further analysis, we ensured that infants were gazing at the fixation object prior to the target object. During the joint attention or no joint attention sequence, the screen was blank for 5000 ms. If the infant became fussy or did not look at the screen, the experimenter gave the infant a short break. The session ended when the infant's attention could no longer be attracted to the screen or if it was clear that the infant was not engaging with the experimenter's face. EEG was recorded continuously and the behaviour of the infants was also video-recorded throughout the session for offline coding of looking behaviour to the experimenter and the screen. Infants' electrical brain activity was measured as they viewed the novel objects presented on the computer screen.

EEG recording and analysis

EEG was recorded continuously with Ag-AgCl electrodes from 19 scalp locations of the 10–20 system, referenced to the vertex (Cz).

FIG. 1. Display of two sequences in the joint attention condition. The adult gazed at the infant's face (top left) and then to a novel object that was displayed on a computer screen for one second (bottom left), and the infant engages with experimenter's face (top right), then attends to the object on the screen (bottom right).

FIG. 2. Display of two sequences in the non-joint attention condition. The adult gazed only at the novel object displayed on the computer screen, with the position of their gaze fixed towards the monitor (top and bottom left); the infant attends to the adult's face (top right) then to the object on the screen (bottom right).

Data was amplified via a Twente Medical Systems 32-channel REFA amplifier and analysed via in-house software. Horizontal and vertical electrooculargram were recorded bipolarly. Sampling rate was set at 250 Hz, EEG data was re-referenced offline to the linked mastoids.

The EEG recordings were segmented into epochs of waveform that comprised a 100-ms baseline featuring a triangular central fixation object and 1000 ms of object presentation. For the elimination of electrical artifacts caused by eye and body movements, EEG data was rejected off-line whenever the standard deviation within a 200-ms gliding window exceeded 80 μV at any electrode. Data were also visually edited offline for artefacts.

Results

Behavioural results

Infant gaze following occurred in 63.7 ± 15.1% (mean ± SD throughout) of joint attention trials and 58.8 ± 16.5% of no joint attention trials. A paired samples *t*-test indicated that there was no significant difference in gaze following between the two conditions. This was also reflected in the mean number of trials used for ERP averaging for the two conditions: joint attention, 11.7 ± 1.54; no joint attention, 11.4 ± 1.5. These gaze-following results are consistent with those reported by Striano & Stahl (2005).

ERP results

For statistical analysis of the Nc component a time window was chosen from 350 to 550 ms after stimulus onset. To assess lateralization, variances of ERPs were analysed by a 2 × 3 repeated measures ANOVA, assessing peak amplitude of the effect. Analysed factors were (i) condition (joint attention × no joint attention) and (ii) lateralization (left × central × right), with the following regions of interest: left fronto-central (F3, FC3, C3), central electrodes (Fz, Cz) and right fronto-central (F4, FC4, C4). We chose these electrodes as in the grand average and in individual averages, the effect was most evident across frontal and central sites (see Fig. 3).

We also assessed for differences in peak latency of the effect in the time window 350–550 ms, utilizing the same analysed factors as reported above.

The ANOVA indicated an effect of location ($F_{1,14} = 8.13$, $P = 0.002$), indicating greater negative amplitude in central locations (−39.86 ± 16.2 μV) than in the left (−32.77 ± 12.03 μV) or the right channels (−30.39 ± 12.11 μV).

There was also an effect of condition by location ($F_{2,28} = 3.69$, $P = 0.038$). A further assessment, utilizing paired sample *t*-tests within each channel group indicated greater negative amplitude to the joint attention condition (−45.85 ± 18.84 μV) relative to the no joint attention condition (−33.87 ± 13.55 μV) in the central electrodes (t_{14} 2.38, $P = 0.032$). No differences were found in left (t_{14} 1.85, $P = 0.086$) or right hemisphere electrodes (t_{14} 0.53, $P = 0.607$). For an example of the Nc effect at these assessed locations, see Fig. 4. For an example of the Nc effect in the context of all scalp locations, see Fig. 3. There were no effects of amplitude latency.

Other effects

Visual inspection of the ERP suggested an early effect in frontal and central leads at ∼150–250 ms after stimulus onset (see Figs 3 and 4), referred to in previous studies as the Pb (reviewed in Webb *et al*.,

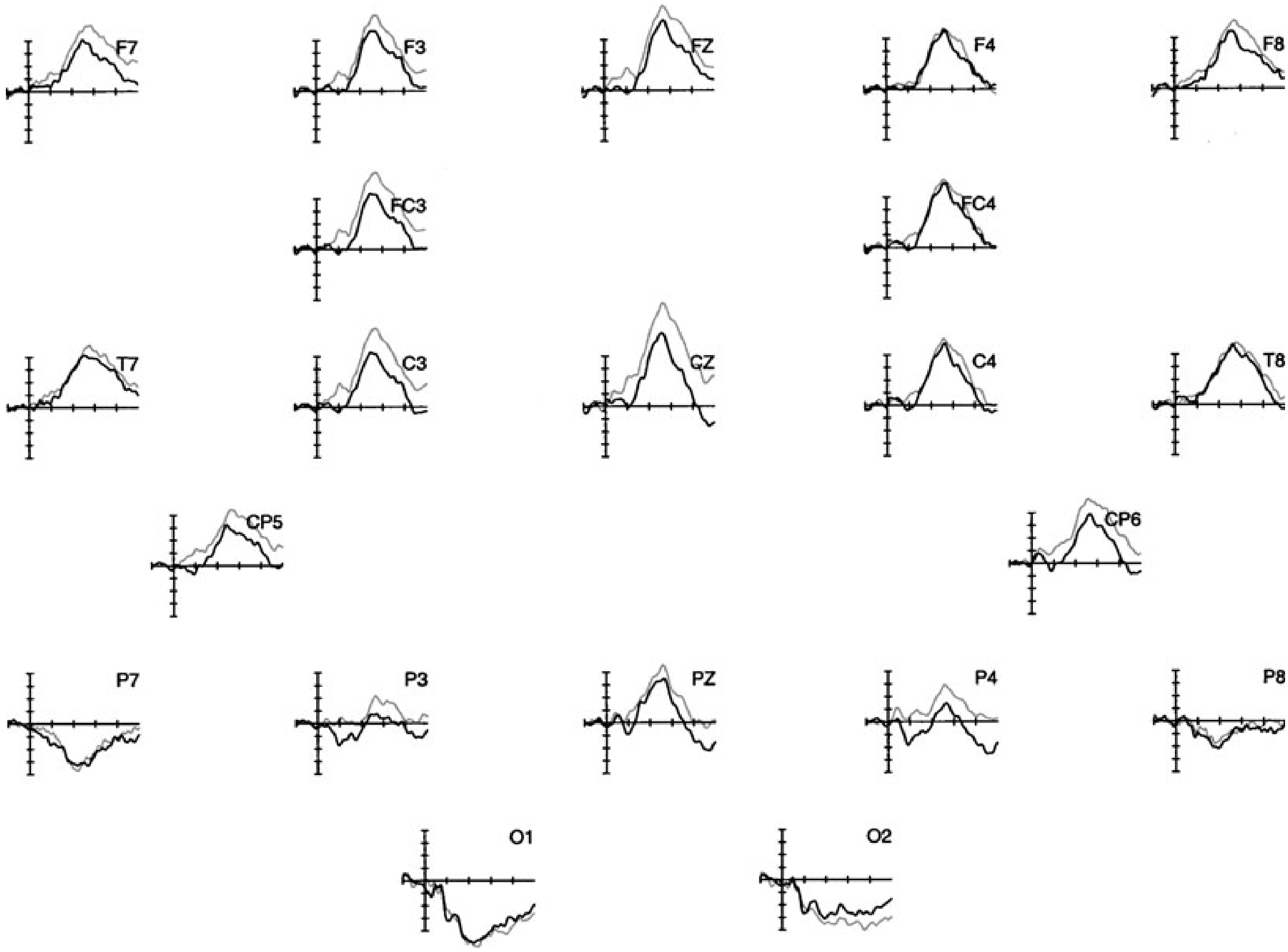

FIG. 3. Grand average ERP signal displaying all recorded electrodes. The Nc effect is most prominent at frontal and central sites. Grey, joint attention condition; black, non-joint attention condition. Horizontal tick mark, 0.2 ms; vertical tick mark, 5 μV.

2005). We utilized the same statistical parameters and electrodes as for our Nc analysis, however, due to the morphology of the waveform, we assessed mean amplitude of this component. The 2 × 3 ANOVA indicated a main effect of condition ($F_{1,14} = 4.65$, $P = 0.049$), reflecting greater amplitude in the joint attention condition (−6.02 ± 11.46 μV) than the no joint attention condition (0.22 ± 8.16 μV). No other significant effects were found.

Discussion

We investigated the negative component (Nc) of the event-related potential (ERP), a well established index of infant attention allocation, in relation to infants viewing objects following interactive joint attention and non-joint attention interaction. Our results indicated a larger peak amplitude of the Nc when infants viewed objects following interactive joint attention relative to the peak amplitude of the Nc following non-joint attention interaction. This result suggests that infants increase attention to aspects of the environment that are more salient. Further research is needed to determine which cues are necessary to establish and maintain joint attention.

The Nc was substantially larger in this study than that seen in previous literature (e.g. de Haan & Nelson, 1997, 2001). This is likely due to the application of the novel paradigm employing live interaction prior to the presentation of the objects on the computer screen. This interactive paradigm most likely provided an increased

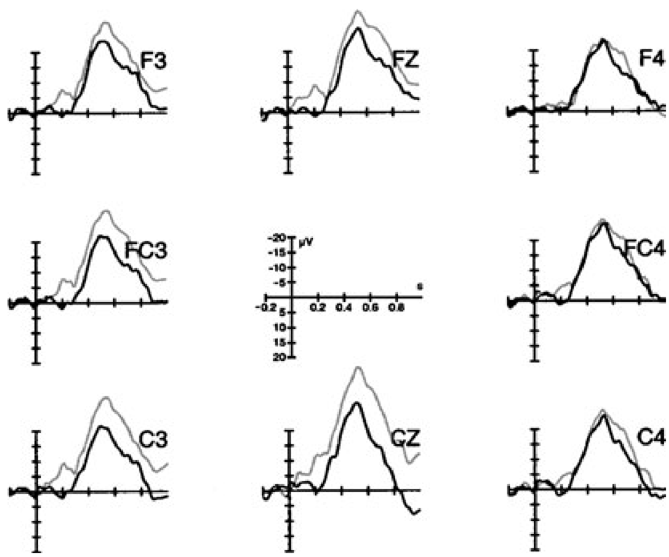

FIG. 4. Grand average ERP signal displaying Nc over fronto-central channels. The effect is evident, particularly at midline sites (Cz, Fz), with a larger amplitude Nc for the joint attention condition (grey) when compared with the non-joint attention condition (black).

social significance during the testing situation for the infant when compared with standard ERP paradigms, which typically feature the passive viewing of stimuli presented on a computer screen. Further support for this interpretation derives from the significant Pb effect found in this study. The Pb component in previous research has been suggested to reflect contextual processing of the objects (see Webb *et al*., 2005). In the present study, the joint attention condition yielded a more negative Pb than the no joint attention condition. The present study suggests a relationship between the Pb and the Nc as the effects were found on the same electrodes with greater negativity in the joint attention condition. However, if a relationship exists, the parameters of any link are currently unclear.

It is possible that the large dropout rate in terms of number of subjects tested and not included means that a specific temperamental type of infant was included in the present investigation. Fifteen infants were included in the final analysis and 52 infants were tested but not included. Potentially those with a larger than mean attention span were included as they could successfully complete the minimum number of trials for criteria for inclusion in the final sample. This possibility cannot be ruled out at this stage, and further work will determine if this is the case. Similarly, as it is well established that 9-month-old infants can engage in joint attention situations, future research is also needed to assess the development of this effect in younger infants.

It is important to note here that the adult's initial eye contact prior to looking at the object was the critical cue to indicate joint attention for infants. Eye contact is generally a cue to joint attention that is readily detected by young infants (Striano & Stahl, 2005). In addition, at least by 4 months of age, eye cuing directed toward objects facilitates neural processing of these objects (Reid *et al*., 2004; Reid & Striano, 2005). The presence of eye-to-eye contact was the critical difference between our joint attention and non-joint attention conditions, but eye contact alone is generally not enough to establish joint attention and therefore probably not enough to faciliate object processing. In support of this view, 4-month-old infants show enhanced cortical processing of faces with direct compared to averted gaze, but only when faces are upright – not when inverted (Farroni *et al*., 2004). Thus, eye contact must be provided in the context of relevant social interaction. In future studies, it will be important to control for eye contact and manipulate other aspects of joint attention, for example the timing of social cues, movement cues, and infant directed speech. Based on prior research (Striano & Stahl, 2005) we predict that the absence of any of these social cues will affect the establishment of joint attention and this will be reflected in the Nc component of the event-related potential. Future studies will be necessary to put these hypotheses to the test. In the current study, we show that eye-to-eye contact, just one important aspect of joint attention, influences neural processing of objects.

In general, joint attention should afford early learning by constraining the allocation of attentional resources when new information is encountered (Gergely & Csibra, 2005). Here we show a larger Nc when infants process objects in joint attention compared to non-joint attention contexts, suggesting that the Nc is one of the neural correlates of joint attention. Given the interplay between joint attention and later cognitive skills such as language development (Baldwin, 1993), the current experiment suggests a general learning mechanism that similarly underlies a range of major cognitive developments across domains. The finding that Nc amplitude is augmented in joint attention contexts is a critical first step in addressing how infants most effectively process and learn new information.

Acknowledgements

The research was supported by the Sofja Kovalevskaja Award granted by the Alexander von Humboldt Foundation, donated by the Federal Ministry of Education, to T. Striano. We are grateful to the infants and partners who participated, to Katharina Kaduk for assistance with testing, and to the Universitäts-frauenklinik and the Eitingon Krankenhaus for assistance with recruitment.

Abbreviations

ERP, event-related potential; Nc, negative component.

References

Baldwin, D.A. (1993) Early referential understanding: infants' ability to recognize referential acts for what they are. *Dev. Psychol.*, **29**, 832–843.

Courchesne, E., Ganz, L. & Norcia, A.M. (1981) Event-related brain potentials to human faces in infants. *Ch. Dev.*, **52**, 804–811.

Farroni, T., Csibra, G., Simion, F. & Johnson, M.H. (2002) Eye contact detection in humans from birth. *Proc. Natl Acad. Sci.*, **99**, 9602–9605.

Farroni, T., Johnson, M.H. & Csibra, G. (2004) Mechanisms of eye gaze perception during infancy. *J. Cogn. Neurosci.*, **16**, 1320–1326.

Gergely, G. & Csibra, G. (2005) The social construction of the cultural mind: Imitative learning as a mechanism of human pedagogy. *Interact. Stud.*, **6**, 463–481.

de Haan, M. & Nelson, C.A. (1997) Recognition of the mothers face by 6-month-old infants. A neurobehavioral study. *Ch. Dev.*, **68**, 187–210.

de Haan, M. & Nelson, C.A. (2001) Electrocortical correlates of face and object recognition by 6-month-old infants. *Dev. Psychol.*, **35**, 1113–1121.

Johnson, M.H. (2005) *Developmental Cognitive Neuroscience*, 2nd Edn. Blackwell, Oxford.

Johnson, M.H. & Morton, J. (1991) *Biology and Cognitive Development: the Case of Face Recognition*. Blackwell, Oxford.

Nelson, C.A. (1994) Neural correlates of recognition memory in the first postnatal year of life. In Dawson, G. & Fischer, K., (Eds), *Human Behavior and the Developing Brain*. Guilford Press, New York, pp. 269–313.

Reid, V.M. & Striano, T. (2005) Adult gaze influences infant attention and object processing: Implications for cognitive neuroscience. *Eur. J. Neurosci.*, **21**, 1763–1766.

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

Reynolds, G.D. & Richards, J.E. (2005) Familiarization, attention, and recognition memory in infancy: an ERP and cortical source localization study. *Dev. Psychol.*, **41**, 598–615.

Richards, J.E. (2003a) Attention affects the recognition of briefly presented visual stimuli in infants: An ERP study. *Dev. Sci.*, **6**, 312–328.

Richards, J.E. (2003b) Development of attentional systems. In de Haan, M. & Johnson, M.H., (Eds), *The Cognitive Neuroscience of Development*. Psychology Press, East Sussex, UK, pp. 73–98.

Rugg, M.D. & Coles, M.G.H. (1995) *Electrophysiology of Mind: Event-Related Brain Potentials and Cognition*. Oxford University Press, London.

Striano, T. & Bertin, E. (2005) Relation among joint attention skills in 5- to 10-month-old infants. *Br. J. Dev. Psychol.*, **23**, 559–568.

Striano, T. & Stahl, D. (2005) Sensitivity to triadic attention in early infancy. *Dev. Sci.*, **4**, 333–343.

Webb, S.J., Long, J.D. & Nelson, C.A. (2005) A longitudinal investigation of visual event-related potentials in the first year of life. *Dev. Sci.*, **8**, 605–616.