An investigation into the mechanisms of inter-brain synchrony during early social interactions

Ira John Marriott Haresign, BSc

School of Psychology, University of East London

Thesis submitted for the degree of Doctor of Philosophy

January 2023

Abstract

Over the last 20 years there has been a growing increase in the amount of research investigating how and why two or more individual's brain activity can synchronise during social interaction. What we know so far from this research is that inter-brain synchrony (defined through temporally coordinated patterns of brain activity between two interacting individuals, Holroyd 2022) tends to associate with moments of behavioural coordination (i.e., when two individuals are doing or attending to the same thing at the same time) and task cooperation (i.e., the action or process of two individuals working together to the same end). These observations have led many researchers to theorise over whether and how behavioural coordination mechanistically drives inter-brain synchrony (Wass et al., 2020; Hamilton, 2021). There is also some very recent evidence to suggest that increased inter-brain synchrony actually facilitates/ supports aspects of social interaction. For example, inter-brain synchrony has been shown to predict team performance (Reinero et al., 2021), although this research is primarily based on correlational study designs.

Taken together however the field of inter-brain synchrony shares one fundamental limitation; that is that it does not account (although see recent animal research e.g., Kingsbury et al., 2019; Zhang et al., 2019), empirically for the mechanisms that give rise to inter-brain synchrony, which would help to falsify claims that inter-brain synchrony is a core mechanism facilitating social interaction. This is because of two main reasons; Firstly, the study of inter-brain synchrony has primarily been investigated as a time-invariant property, almost no studies have explored how inter-brain synchrony varies over time relative to individual moments of behavioural coordination. Secondly, little attention has been paid to the changes

in the underlying signal properties (i.e., increases in power, changes in frequency) that must take place for two unsynchronised signals to become synchronised (e.g., Haresign et al., 2022).

Using two-person naturalistic biobehavioural recording techniques, coupled with state of the art, EEG pre-processing and analyses procedures (see chapters 5 and 6), the present thesis examines the mechanisms that give rise to inter-brain synchrony during parent-infant social interactions.

Evidence is presented showing how inter-brain synchrony does not arise around individual moments of gaze coordination. This is despite previous investigations suggesting that increased inter-brain synchrony (averaged over all moments of eye contact) associates with gaze synchrony. Evidence also shows the contribution of behavioural coordination across multiple modalities to inter-brain synchrony during parent-infant social interaction

Discussion is focused on the contribution of these findings to our understanding of the mechanisms that give rise to inter-brain synchrony.

Keywords: EEG, hyperscanning, inter-brain synchrony, social interaction, early development

Declaration

This thesis comprised data from one study. The empirical chapters (chapters five, six, seven and eight) used data from the Leverhulme project, headed by Sam Wass and funded by the Leverhulme trust. Data collection was shared between the three team members of the Leverhulme project: myself, Megan Whitehorn and Emily Phillips. Due to the overlapping nature of our intervention design we each took responsibility for collecting data for different aspects of the study (pre/post data and intervention data). I ran data collection for the pre and the post intervention sessions with one other member of the team, in which we collected naturalistic dual EEG and ECG recordings. The setup up of the dual EEG/ ECG recording system and the development of the simulations/ interventions used during data collection was a team effort alongside Sam Wass. Participant recruitment, data management and security, as well as the training of research assistants was shared between all members of the team. As was the application for ethics. The present thesis represents my own work.

Of the four empirical chapters of my PhD, three are based on shared data. As such some of the analysis in these chapters will reflect primary analyses, and some will reflect secondary analyses.

For the first author publications which form the basis of the work presented in the empirical chapters, all of the analyses, the research questions, analytical plan, statistical scripts, and results interpretations are my own work, under supervision of Sam Wass. All first author publications entail original data collection and analyses.

Distinct contribution

(1) *Naturalistic methodologies in parent-infant interaction*. Whereas previous studies investigating the neural correlates of infant's social attention have focused on static screen-based stimuli, the work presented in the empirical chapters measures brain activity during free-flowing naturalistic interactions.

(2) *The use of dual EEG or 'hyper scanning'.* Work included in the empirical chapters will also all revolve around the use of dual EEG. Very few studies have used such techniques to study parent infants' interactions.

(3) *The advancement of methods associated with measurement of inter brain synchrony between parents and infants.* Chapter six aims to address issues surrounding the quantification of inter-brain synchrony between infants and adults. To date there exists no comprehensive guidelines for estimating inter brain phase synchrony in such populations.

COVID-19 impact statement

The work for this thesis took place between January 2019 and January 2023, therefore it is worth noting that a large amount of the data collection took place during the COVID-19 pandemic. Prior to the first lockdown of 2020 we had already collected 48 full datasets. Fortunately, we were able to resume testing after just a few months of lockdown and began running sessions with families again in the summer of 2020. Additional risk assessments (see appendix B) were completed by the research team and approved by UEL; to mitigate the risks of COVID-19 transmission participants were instructed to fill out all consent forms online prior to the session. During the sessions researchers and participants wore full PPE (face masks and gloves). Sessions took place in a well-ventilated room. After each session every surface and toy were then cleaned and disinfected.

The analyses presented in this thesis reflect a combination of data that was collected before, during and after the pandemic. However, it is beyond the scope of this thesis to examine the impact of the COVID-19 pandemic on parent or infant behaviour or neural activity. In my analyses I have assumed that the core neural mechanisms under investigation were unchanged. It is my aim to make our data publicly available after the study is completed in order to facilitate any investigation into the impact of the COVID-19 pandemic on early development and or parent-infant social interactions.

List of original publications

First author publications

1. Haresign, I. M., Phillips, E., Whitehorn, M., Noreika, V., Jones, E. J. H., Leong, V., & Wass, S. V. (2021). Automatic classification of ICA components from infant EEG using MARA. *Developmental cognitive neuroscience*, *52*, 101024 (*). https://doi.org/10.1016/j.dcn.2021.101024

In publication 1 I collected the data, alongside the other PhD students. I performed the analyses and wrote the manuscript. The work from publication 1 forms the basis of chapter 5 of the present thesis.

2. Haresign, I. M., Phillips, E. A. M., Whitehorn, M., Goupil, L., Noreika, V., Leong, V., & Wass, S. V. (2022). Measuring the temporal dynamics of inter-personal neural entrainment in continuous child-adult EEG hyperscanning data. *Developmental cognitive neuroscience*, *54*, 101093 (*). https://doi.org/10.1016/j.dcn.2022.101093

In publication 2 I simulated the data that was used, I performed the analyses and wrote the manuscript. The work from publication 2 forms the basis of chapter 6 of the present thesis.

3. Haresign, I. M., Phillips, E. A. M., Whitehorn, M., Lamagna, F., Eliano, M., Goupil, L., & Wass, S. V. (2022). Gaze onsets during naturalistic infant-caregiver interaction associate with 'sender' but not 'receiver 'neural responses, and do not lead to changes in inter-brain

synchrony. *Scientific reports*, *In press* (*). Pre-print available online: https://doi.org/10.1101/2022.05.27.493545

In publication 3 I collected the data, alongside the other PhD students. I performed the analyses and wrote the manuscript. The work from publication 3 forms the basis of chapter 7 of the present thesis.

Co-author publications

4. Wass, S. V., Whitehorn, M., Haresign, I. M., Phillips, E., & Leong, V. (2020). Interpersonal neural entrainment during early social interaction. *Trends in cognitive sciences*, *24*(4), 329-342. https://doi.org/10.1016/j.tics.2020.01.006

In publication 4 I contributed to discussion that formed the basis of the ideas presented within the review. I also helped with editing the manuscript. Many of the ideas discussed within publication 4 form the basis of the chapter 2 of the present thesis.

5. Kayhan, E., Matthes, D., Haresign, I. M., Bánki, A., Michel, C., Langeloh, M., & Hoehl, S. (2022). DEEP: A dual EEG pipeline for developmental hyperscanning studies. *Developmental cognitive neuroscience*, *54*, 101104. https://doi.org/10.1016/j.dcn.2022.101104

In publication 5 I contributed code for the pre-processing of adult-infant EEG hyperscanning data that was published as part of publication 5. I also helped with editing the manuscript. This work did not contribute directly to any chapter within the present thesis.

6. Phillips, E., Goupil, L., Haresign, I. M., Bruce-Gardyne, E., Csolsim, F. A., Whitehorn, M., & Wass, S. (2021). Proactive or reactive? Neural oscillatory insight into the leader-follower dynamics of early infant-caregiver interaction. *Proceedings of the National Academy of Sciences, in press (*)*

In publication 6 I contributed to the data collection. I also helped with editing the manuscript. This work did not contribute directly to any chapter within the present thesis.

7. Daubney, K., Suata, Z., Haresign, I. M., Thomas, M., Kushnerenko, E., & Wass, S. V. (2022). The development of the relationship between auditory and visual neural sensitivity and autonomic arousal from 6m to 12m. *bioRxiv*. Pre-print available online: https://doi.org/10.1101/2022.12.15.520575

In publication 7 I contributed to the analysis of the data. I performed several of the ERP analyses that are presented within the paper. I also helped with editing the manuscript. This work did not contribute directly to any chapter within the present thesis.

8. Phillips, E., Goupil, L., Whitehorn, M., Bruce-Gardyne, E., Csolsim, F. A., Kaur, N., Haresign, I. M., & Wass, S. (2023). Endogenous oscillatory rhythms and interactive contingencies jointly influence infant attention during early infant-caregiver interaction, *in prep (*)*

In publication 8 I contributed to the data collection. I also helped with editing the manuscript. This work did not contribute directly to any chapter within the present thesis.

Table of Contents

Table of Figures

Table of Tables

Definition of key terms

Synchrony. 'Coordination of biological and social processes during social contact' (Feldman, 2015, p. 369); in dyadic contexts, synchrony may be operationalised as 'concurrent' ('when A is high, B is high') or 'sequential' ('changes in A forward-predict changes in B'; Helm et al., 2018; Wass et al., 2020).

Entrainment. Entrainment is defined by a temporal locking process in which one system's motion or signal periodicity becomes temporally aligned with the periodicity of another system. This process is a universal phenomenon that can be observed in physical (e.g., pendulum clocks) and biological systems (e.g., fireflies) (Thuat et al., 2015).

Phase resetting. An abrupt shift in oscillatory phase. Often associated with the onset of some stimulus/ event.

Phase locking. Defined as temporal alignment of phase between two signals, wherein two signals show consistent variation in phase with respect to one another.

Hyperscanning. The recording of brain activity from multiple individuals

Abbreviations

Used in the synopsis or in the original publications.

AOI	area of interest
ANOVA	analysis of variance
EEG	electroencephalogram
EMG	electromyogram
ECG	electrocardiogram
EOG	electrooculogram
ERP	event-related potential
ET	eye-tracker, eye-tracking
MSE	mean squared error
PLV	phase locking value or “ phase locking”
PDC	partially directed coherence
PFC	pre-frontal cortex

Acknowledgements

Firstly, I would like to thank all of the families that gave up many hours of their time to make this research possible. Funding was provided by the Leverhulme trust and institutional support was provided by the University of East London.

I would like to thank Sam Wass for the incredible amount of support and guidance he has given me during my PhD. Throughout the process Sam's attitude has been caring and patient and his passion for science has been a continual source of inspiration and motivation. I would also like to thank Victoria Leong whose technical knowledge was a great recourse. My sincerest gratitude also goes out to all of the collaborators that I have worked with over the last four years; Louise Goupil and Emily Jones in particular who I have learnt a great deal from.

It has been a great pleasure to work with and get to know Emily Phillips and Megan Whitehorn. Certainly, this has not been a process without challenge, but I sincerely hope that they will remember my involvement in their journeys as fondly as I will remember their involvement in mine. I am also grateful for every member of the UEL babylab that contribute to making it such an enjoyable place to work.

Lastly, I'd like to also extend a special thank you to Angela Gosling whose belief in me at a very early stage made this entire PhD possible.

Chapter 1 – Introduction

The ability to coordinate my behaviour with those of a social partner is seen throughout the natural world across multiple species. Yet, it is remarkably complex. To coordinate my actions with someone else I must track my own internal states and behaviours and, simultaneously, my partner's internal states and behaviours, and I must do so in a time-sensitive and context-dependent manner.

Previous research has for decades documented how early parent-infant social interactions give rise to synchronous (temporally coordinated) patterns of behaviour and the importance of this for later development (Feldman, 2007). For example, synchrony of behaviour (e.g., eye gaze) during early infant-caregiver interactions is thought to be crucial for developmental outcomes including self-regulation (Feldman & Greenbaum, 1997), IQ (Feldman & Greenbaum, 1997) and language development (Rogoff, 1990). Further research has suggested behavioural synchrony during early social interactions is important as it presents opportunities for increased information exchange across the dyad (Feldman, 2007; Leong et al., 2017) as well as possibly providing temporal structure to the interaction, which could ensure that relevant information is received/ processed during times when infants are maximally sensitive (Wass et al., 2020). Yet although, much is known about the importance of behavioural synchrony during early social interactions, remarkably little is known about how the developing brain substantiates this.

Currently almost everything we understand about how the infant brain supports social interaction comes from studies that have used screen-based experiments, in which brain

activity is averaged relative to the repeated presentation of a social stimulus. This approach has taught us a lot. For example, from this research we know that even from very early on in development our brains are highly sensitive to ostensive cues; Even from just a few months of life infant's brains show differential neural responses to images faces showing direct vs averted eye contact (Farroni et al., 2002). Further during live social interaction, activity in the frontal cortex of infants increases during direct gaze compared with averted gaze (Grossmann et al., 2008; Urakawa, et al, 2015). Additionally, the infant brain is more sensitive to infant-directed speech compared to adult-directed speech (Zhang et al., 2011), and during live interactions, the infant prefrontal cortex tracks and responds to increases in the infant-directedness of speech (Piazza, et al, 2019). Moreover, although most studies have only used static images of faces, presented on a screen, some other studies have explored infants' sensitivity to ostensive cues within triadic attention situations (i.e., the sharing of attention between child, social partner, and object). These studies have suggested that, when a live adult gazes first to an infants' face and then to an object, evoked neural responses were increased (Striano et al., 2006; Hoehl et al., 2014).

However, we know that the neural mechanisms involved in these screen-based simulations are not the same as those involved in real social interactions. For example, in adults, larger amplitude ERP responses are observed across occipital electrodes in response to live eye contact with another person compared to the screen-based presentation of face showing eye contact (Ponkanen et al., 2011). A further example of this is reflected in the audience effect (Hamilton & Lind, 2016; see also Cañigueral et al., 2022 for a discussion), which is the effect of the physical presence of another person(s) who is watching the participant causing changes in the participant's behaviour (e.g., Cañigueral & Hamilton, 2019b) and brain activity (e.g., Izuma et al., 2010a). Therefore, as many authors have recently emphasised (Schilbach et al.,

2013; Redcay & Shilbach, 2019; Wass et al., 2020; Wass & Goupil, 2022; Cañigueral et al., 2022) a full understanding of the social brain requires investigations that consider the complexity and dynamic nature of social interactions as well as the interpersonal influences on brain function that arise as a product of this.

The two-person neuroscience of naturalistic social interactions

Only very recently due to technological advances, allowing for the simultaneous measurement of brain activity from two interacting individuals (hyperscanning) has it become possible to examine how social interactions are substantiated across two interacting brains. One way to investigate this is to examine how patterns of individual brain activity are influenced by the live social context. For example, Menoret and colleagues investigated neural markers of goal representation during a joint action task using dual EEG. Here the observers watched actions performed by either a human actor and or by a robot. The interactional context of these actions was also manipulated depending on whether the observer was required to act contingently or do nothing. The analyses revealed a concomitant suppression of beta oscillations in both the actor's and the observer's EEG, time locked to the onset of the actor's movement. This suppression was found to be greater during interactive vs non interactive conditions and irrespective of whether the actor was a human or a robot, suggesting that the social interaction context differentially activation motor brain regions. Additionally, previous research with adults has shown that joint action planning resulted in amplitude modulations of the frontal ERPs compared with individual action planning (Kourtis et a., 2013). Similar ERP components have also been previously associated with decision making and updating of representations.

Inter-brain synchrony during social interactions

An additional approach is to examine how shared patterns of brain activity across individuals support social interactions; At the neural level, inter-brain synchrony can be defined as a dyadic mechanism, wherein temporally coordinated patterns of brain activity between two interacting individuals supports aspects of their ongoing social interaction (Holroyd 2022). A growing body of research is now examining how inter-brain synchrony supports social interaction. Many empirical studies have examined under what circumstances of relative to which behaviours is inter-brain synchrony enhanced. For example, a number of studies have observed increased inter-brain synchrony during mutual gaze. The majority of this research claims to measure inter-brain synchrony, although we recognise that not all of these studies will meet the framework of inter-brain synchrony set out in more recent theoretical accounts (Holroyd, 2022). Kinreich and colleagues (2017) observed significantly correlated gamma (30-60Hz) activity between interacting adults during social interaction. Higher interpersonal gamma correlations were also associated more strongly with mutual vs non-mutual gaze. Similarly, Luft and colleagues (2021) found that mutual gaze was associated with higher inter-brain gamma band (30-45Hz) coherence (a spectral measure based on correlation) between interacting adults than non-mutual gaze. In the developmental literature, our group investigated inter-brain synchrony in 7.5-month infant-adult dyads during moments of mutual and non-mutual gaze (Leong et al., 2017). During a live social, but not interactional condition infants observed an adult singing nursery rhymes, who was instructed to look either directly at the infant, directly at the infant with their head turned at a slight angle, or away from the infant. Consistent with research on adults, we found greater infant-adult neural synchrony during moments of mutual vs non-mutual gaze, measured using partially directed coherence (PDC-a spectral Granger causal measure of synchrony) in Theta (3-6Hz) and Alpha (6-9Hz)

band activity. This study thus suggests that the impact of mutual gaze on inter-brain synchrony found in adult-adult dyads (Kinreich et al., 2017; Luft et al., 2021) is already present early on in development, though possibly in lower frequency brain rhythms. Additionally, associations between mutual gaze and adult-infant inter-brain synchrony have been found using fNIRS (Piazza et al., 2020). However, because the temporal resolution is so different between EEG and fNIRS it is difficult to conclude that these represent the same neural mechanisms.

Despite this currently we understand virtually nothing about how inter-brain synchrony arises during social interactions. This is primarily because of an over emphasis on non-event locked approaches. Here, using a sliding window approach, inter-brain synchrony is calculated within a short segment of the data (typically 1-2 seconds in EEG data (Leong et al., 2017; Santamaria et al., 2020) and around ~30 seconds in fNIRS data (Nyguyen et al., 2021)). This window is then shifted forwards in time by a given increment (typically half the number of time points within the window) and inter-brain synchrony is calculated again. This yields a single value per segment which is then averaged across all segments to a single value representing the amount of inter-brain synchrony throughout the entire continuous stream of data. The amount of inter-brain synchrony is then typically compared between experimental conditions (e.g., Leong et al., 2017; Nguyen et al., 2021; Piazza et al., 2020; Santamaria et al., 2020), and/or is correlated with an outcome measure (Davidesco et al., 2023; Nguyen et al., 2020).

This approach has two fundamental limitations; First, it will never allow researchers to be sure that observed effects are not attributable to underlying artifact. For example, during social interaction eye movements are not random but are influenced a partner's behaviour

(Valtakari et al., 2021). For example, Figure 1a shows adults' gaze behaviour time-locked to the moment where the infant looks up to the adult during a face-to-face interaction. An increase in the likelihood of the adult looking to the infant is observed in the 1000msec window following the infant gaze shift. Based on the argument given above, we know that this will lead (not in every case, but often), to eye movement artifact in the infant's data, followed by eye movement artifact in the adult's data shortly afterwards. Figures 1b-1i, which are based on simulated data, illustrate how this pattern can give rise to the impression of inter-brain synchrony. Artifact manifests both in the time and frequency domains. Because of the non-event locked analyses do no examine temporal dynamics of inter-brain synchrony relative to behavioural events they will never be able to fully understand the behavioural dynamics. Further because these approaches often involved heavy down sampling of the data prior to/ during the calculation of inter-brain synchrony (need examples here), which can create the impression that two events that occur slightly after one another in time are actually occurring at the same time, resulting in concurrent inter-brain synchrony. Additionally, this approach will never allow researchers to understand the mechanisms that give rise to associations in brain activity between two interacting individuals.

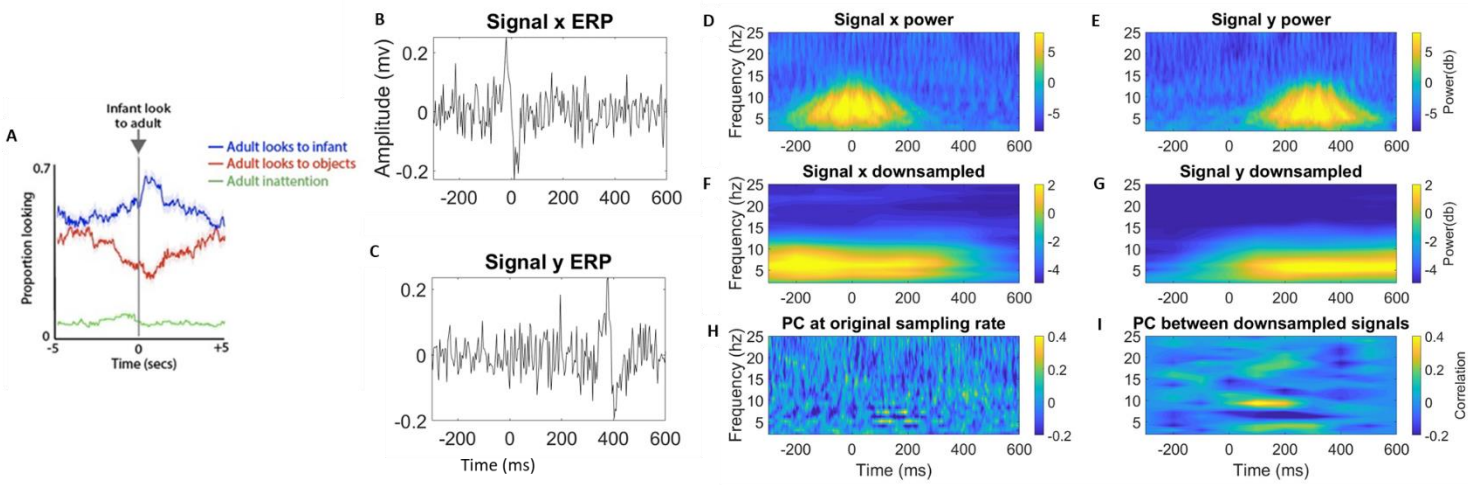


Figure 1.1. Illustration that synchrony during social interaction often results from lagged contingent responding and neural activity both precedes and follows these behaviours. A) Probability of changes in adults gaze peaks around ~1 second after changes in infant gaze. B) *ERP plot of signal 'x'. C) ERP plot of signal 'y'. D) Time-frequency power of signal x. E) Time-frequency power of signal y. F) Down sampled, time-frequency power of signal x. G) Down sampled, time-frequency power of signal y. H) Spearman's correlation of single trial power (PC) between x and y) computed at each time-frequency point (i.e., original temporal scale of data). I) Spearman's correlation of single trial power between x and y computed on the down sampled data. The AOIs on panel f indicate regions of significant correlations.*

Mechanisms for interpersonal neural synchrony during social interactions

One candidate mechanism for how inter-brain synchrony is established/ maintained is through phase resetting (i.e., an abrupt change or shift in phase of an oscillation) of neuronal oscillations. The inter-brain phase reset theory asserts that certain behavioural events may drive simultaneous changes in phase in both interacting brains (Wass et al, 2020; Leong et al., 2017). It is known that the phase of neuronal oscillations reflects the excitability of underlying neuronal populations to incoming sensory stimulation (Klimesch et al., 2007; Jensen et al., 2014) and that sensory information arriving during high-receptivity periods is more likely to be perceived than information arriving during low-receptivity periods (Busch et al., 2009; Mathewson et al., 2009; 2010; 2011; 2012). This suggests that there is an optimal (range of) phase for perceiving information. Therefore, it has been argued that there must be mechanisms (phase resetting) for modulating the phase of neuronal oscillations, in order to match the temporal structure of the environmental input (van Diepen et al., 2015; Ruzzoli et al., 2019), enabling enhanced processing. Empirical evidence supports the role of phase

resetting as an intra-brain mechanism, facilitating neural entrainment to temporal structures within the environment, for example speech (Giraud & Poepell 2012; Biau et al., 2015). It is therefore logical to question whether similar mechanisms also operate at the interpersonal level. For example, inter-brain synchronisation may increase within a dyad following the onset of communicative signals (such as gaze, gestures, or vocalisations) that reset the phase of both interacting partners. Here, neural oscillations in both the sender (of the social signal) and the receiver's brain that were previously random with respect to each other (low inter-brain synchrony) would be simultaneously reset in response to a communicative signal. Following this reset the neural activity of both the sender and the receiver would oscillate with more consistent variation over time (high inter-brain synchrony).

A second possible mechanism is that the interplay between action and prediction involved in turn-taking during social interactions could drive changes in inter-brain synchrony. The key idea here is that if the neural systems responsible for planning (Dikker et al., 2014; Shamay-Tsoory et al., 2019; Hamilton, 2021) and generating (e.g., Hirsch et al., 2017; Kirkland, 2020) one's own behaviours are co-localized with those responsible for predicting a partner's behaviours then at times when both individuals are required to mutually predict each other's behaviours, inter-brain entrainment could increase (Hamilton, 2021). For example, using non-event locked methods of analysis, Kingsbury and colleagues recently provided some support for this hypothesis in interacting mice. They used micro endoscopic calcium imaging (which has a relatively low temporal resolution of ~1sec) to record from the dorsomedial prefrontal cortex (dmPFC) of interacting mice during social interactions (Kingsbury et al., 2019). General linear modelling of each individual's brain activity relative to their own and their partners behaviour, showed that neuronal activity associated with encoding the mouse's own actions was also associated with encoding the actions of the partner mouse. Further

coherence in the activity across the dmPFC of both mice disappeared when these self/other coding neurons were removed (Kingsbury et al., 2019). Here, changes in inter-brain synchrony could be measured concurrently at times when both partners are required to mutually predict an event or behaviour. Prediction would drive changes in either power or phase in both partners (e.g., Mandel et al., 2016; Bögels, 2020), the relationships between the two could be examined using concurrent measures of inter-brain entrainment. Additionally, action-prediction during turn taking could also be examined using measures of sequential entrainment. Here sequential inter-brain entrainment could be measured between changes in power/ phase in the predictor's brain activity supporting prediction of a partner's upcoming behaviour and changes in the actor's brain activity facilitating the action/ behaviour at some time lag later. Here inter-brain entrainment might peak around these 'handover' moments and decrease before and after. This mechanism might also be influenced by factors such as the amount (e.g., Nguyen, et al, 2020) and (perceived) quality (e.g., Bloom, 1988) of turn-taking.

The investigation detailed with the present thesis aimed directly to replicate and extend previous research by supervisors Dr Sam Wass and Dr Victoria Leong, published in *PNAS* in 2017 (Leong et al., 2017). In this study Leong and colleagues measured associations between EEG based, inter-brain synchrony, using granger-causal methods and adult-infant (7.5-month-old) mutual gaze. They used a semi-naturalistic paradigm in which infants observed adults singing nursery rhymes (live). The adult was a-priori instructed to maintain either direct eye contact with the infant, averted (head and gaze) eye contact with the infant, or directed obliques (head, but not gaze averted). Inter-brain synchrony in the theta (3-6Hz) and alpha (6-9Hz) bandwidths was compared between gaze conditions, averaged across the entire interaction. Inter-brain synchrony during the direct and direct oblique gaze conditions were both significantly higher than in the indirect gaze condition. Further, infants vocalized more

frequently during live direct gaze, and individual infants who vocalised longer also elicited stronger synchronisation from the adult.

These results suggest that inter-brain synchrony plays a role in facilitating adult-infant communication during moments of mutual gaze and provide direct support for several highly influential theories within the developmental literature, which suggest that from birth humans are preferentially sensitive to mutual gaze and that this plays a special role within development. In their highly influential paper in 2009, Csibra and Gergely outlined their theory of 'Natural Pedagogy' which proposed that human communication is engineered to facilitate information transfer across the dyad. From birth infants are primed to be on the receptive side of this, by being highly sensitive to ostensive cues (most evidence is eye contact) and that these cues prime the infant brain to be maximally sensitive to receiving new information.

A key question raised by Leong and colleagues' results, however, was how exactly inter-brain synchrony is being established during adult-infant interactions. We know from many studies that have explored face perception using screen-based stimuli that passively receiving information, e.g., when presents a child (Farroni et al., 2002; Taylor et al., 2004; Conte et al., 2020) or an adult (Wantanbe et al., 2002; Conty et al., 2007) with an image of a face, whilst they are continually fixated on the screen, elicits a neural response. Typically, this is measured as a series of ERP components, e.g., P1, N170/N290, P300/P400 for adults/ children respectively. Therefore, one possibility is that the inter-brain synchrony observed in Leong and colleagues' study was simple an alternative way of measuring basic visual processing related neural responses in adults and infant whilst they interact. An alternative possibility is that the observed inter-brain synchrony is measuring some other neural

processes/ cognitive function that is unique to live social interaction and therefore could not have been measured in these previous screen-based studies. This was not addressed within Leong and colleagues' study and was the primary aim of the present thesis.

Overview

The aim of this thesis was to investigate the mechanisms that give rise the inter-brain synchrony. We focused our investigation on early social interactions in which temporally coordinated behaviour is of increased importance. More specifically we were primarily interested in extending Leong and colleagues' results (discussed above), examining how inter-brain synchrony arises around moments of parent-infant mutual gaze. The results of the analyses presented in this thesis provide foundational knowledge about how inter-brain entrainment arises during early social interactions.

Bridging the gap between passive screen-based and active naturalistic social neuroscience however presents several substantial challenges that will need to be considered in any study. Firstly, as postural and eye movements introduce artifacts into EEG recordings, which are often orders of magnitude larger than signals generated from within the brain (Delorme, 2022), the development of social perception and cognition is routinely studied under laboratory conditions that restrict motion. However, this is in contrast to the way that infants typically dynamically and actively explore their environment (Gottlieb et al., 2013; Goupil & Proust, 2023; Kidd et al., 2012; Oudeyer & Smith, 2016; Poli et al., 2020) and 'forage' for information (Robertson et al., 2004). Here information is presented on a screen and only when subjects are still and attentive operationalised through sustained visual fixation. By constricting infants' ability to actively engage with their environment, the cognitive and perceptual processes under investigation may fundamentally differ from those involved in

real-world social interactions, thus results obtained through studies that prioritise real-world validity may come to challenge previous findings from screen-based simulacra (for an example see, Schiller et al., 2004). Secondly, only more recently has research begun to think about how current EEG analysis methodology can be adapted to investigate interpersonal neural influences and coordination and the challenges associated with this.

Alternatively, which is the approach presented within this thesis, researchers could analyse infants' perception and cognition during free-flowing social interactions in the presence of a social partner (which we know impacts brain function, e.g., Izuma et al., 2010a; 2010b; Ponkanen et al., 2011a; 2011b; Wass et al., 2018). Here brain activity is recorded from both interacting individuals and patterns of intra, and inter-brain activity are analysed around moments of naturally occurring behavioural coordination. Over the past decade there has been an increasing number of studies that have taken this approach. These studies have revealed striking patterns of inter-brain synchrony that arise between adult-adult and adult-infant dyads during social interaction. However, a review of this research suggests they largely share one common and fundamental limitation; they do not present empirical evidence for how mechanistically these patterns of inter-brain synchrony emerge. In particular, from a methodological viewpoint what are the necessary changes that must occur within two signal to facilitate this and from a more theoretical stance what aspects of behaviour drive changes in inter-brain synchrony and how is this influenced by other cognitive processes e.g., attention., motivation.

The overarching goal of the present thesis was to explore the mechanisms that might give rise to patterns of temporally coordinated brain activity (termed inter-brain synchrony) during parent-infant social interactions. To facilitate this, we recorded EEG activity from parents and

infants simultaneously (EEG hyperscanning) whilst they engaged in naturalistic, free play interactions. From this we were able to analyse intra and inter brain dynamics from both parents and infants to try to further understand how temporally coordinated patterns of brain activity between parents and infants arise during social interactions.

The first part of this thesis focuses on the substantial number of methodological challenges associated with the preparation and analysis of parent-infant dual EEG data recorded during un-restricted, naturalistic, social interactions. For example, compared to screen-based tasks, in naturalistic paradigms, the 'simulation' can be child-controlled (e.g., the child turning to the parent in a naturalistic interaction), and so artifacts, particular related to eye-movements are more likely to be time-locked to neural signals of interest; the removal of artifact is thus likely to also affect the analysis of neural signals. Thus, we needed to develop approaches (that were tuned specifically to infant EEG data) that remove artifactual signals from the EEG recording without removing entire sections of the data, as is routinely done in screen-based research. Secondly only more recently has research begun to think about how EEG analysis techniques can be adapted to investigate inter-brain synchrony and the challenges associated with this. Therefore, we needed to adapt tools from the literature on analysing intra-brain synchrony to work for inter-brain synchrony analyses and explore how well these were able to capture a given relationship between two individual's patterns of brain activity.

Having established our methodology, the second part of this thesis explores the mechanisms that support the establishment of inter-brain synchrony during social interaction, in which we focus on moments of behavioural coordination extracted from parent infant social interactions, examining how patterns of inter-brain synchrony around these moments propagate through topographical and time-frequency space. Through event-related analyses

of the scalp-recorded EEG we highlight the utility of this approach, by showing how results obtained from naturalistic paradigms closely resemble those previously reported from screen-based research.

Our results challenge the conclusions drawn from both traditional screen-based investigation of eye processing in infancy (e.g., Farroni et al, 2002) . First, when we repeat the Farroni et al analysis, we find no evidence to support the hypothesis that, during real-world naturalistic social exchanges, infant's brains respond differently when someone looks directly at them. This is despite the fact our data are clean, as shown by the clear face/object ERPs that we include in the SM, and the other data quality metrics that we also report in the SM. Our sample size was also larger than in the original study (N=55 vs N=16 in the original study).

Second, when we examine inter-brain entrainment during direct vs averted gaze, we also find no evidence to replicate the previous finding, of Leong and colleagues (2017) that Granger-predictive associations in brain activity are stronger during direct gaze. This is, again, despite that our sample size was larger than in the original study (N=55 vs N=17 per group). Instead, our findings suggest a new conclusion: that the effects of mutual gaze are strongest at the intra-brain level, in the 'sender' but not the 'receiver' of the mutual gaze.

The findings presented in this thesis will be crucial in adding empirical evidence to recent theoretical papers that 'push back' against the rapid growth of interest in dyadic neuroimaging recording techniques that is currently taking place around the world. Given that they go to the very heart of our current prevailing theories of how our brains learn to process social information.

Chapter 2 - Measuring the temporal dynamics of inter-personal neural entrainment in continuous child-adult EEG hyperscanning data

The following chapter is a publication of an original article investigating methods for measuring the temporal dynamics of parent-infant inter-brain synchrony (Marriott Haresign et al., 2022). Subheadings, figure placement, figure and table numbers, and citation style have been adapted to conform to the general thesis format. The supplementary materials (SM) for this publication are also presented within this chapter.

Abstract

Current approaches to analysing EEG hyperscanning data in the developmental literature typically consider interpersonal entrainment between interacting physiological systems as a

time-invariant property. This approach obscures crucial information about how entrainment between interacting systems is established and maintained over time. Here, we describe methods, and present computational algorithms, that will allow researchers to address this gap in the literature. We focus on how two different approaches to measuring entrainment, namely concurrent (e.g., power correlations, phase locking) and sequential (e.g., Granger causality) measures, can be applied to three aspects of the brain signal: amplitude, power, and phase. We guide the reader through worked examples using simulated data on how to leverage these methods to measure changes in interbrain entrainment. For each, we aim to provide a detailed explanation of the interpretation and application of these analyses when studying neural entrainment during early social interactions.

2.1. Introduction

Behavioural evidence suggests that social factors influence how infants pay attention (Yu & Smith, 2016) and learn (Kuhl et al., 2003) during early life. But we currently understand little about how these interpersonal influences are instantiated in the brain (Wass et al., 2020; Redcay & Schilbach, 2019; Redcay & Warnell, 2018; Hoehl et al., 2021). Hyperscanning is a method of simultaneously acquiring neural activity from two or more individuals that allows insights into these questions (Dumas et al., 2010; Schilbach et al., 2013). Hyperscanning approaches are often paralleled with an emphasis on using more free-flowing 'naturalistic' study designs that record brain activity during real-life interactions – rather than studying neural responses to repetitive and unecological, trial-based tasks administered via a computer.

Recently, research with non-human animals (e.g., Kingsbury et al., 2019; Zhang & Yartsev, 2019) and human adults (Liu et al., 2018; Redcay & Schilbach, 2019), as well as research with children/infants using fNIRS (Nguyen et al., 2021; Piazza et al., 2021; Reindl et al., 2018) and electroencephalography (EEG; Leong et al., 2017; Wass et al., 2018) has started to use hyperscanning to uncover complex patterns of interbrain entrainment (IBE) during social interaction. Research relying on EEG with child/infant populations has shown that bidirectional Granger-causal influences between infants' and adults' neural activity are greater in theta (3-6Hz) and alpha frequency bands (6-9Hz) during moments of mutual than non-mutual/ averted gaze (Leong et al., 2017). We also know that patterns of IBE in the theta and alpha bands are higher when adults model positive emotions during social interaction than when adults model negative emotions (Santamaria et al., 2020). These findings suggest that - consistent with fNIRS studies that have shown IBE patterns over longer temporal scales (e.g., Piazza et al., 2021; Nguyen, et al., 2020; 2021) - IBE may also be discernible at the more fine-grained, sub-second scale studied using EEG.

All these approaches used thus far, however, share one fundamental limitation. Hyper-scanning researchers typically calculate the amount of IBE observed between two interacting partners averaged across whole experimental conditions (Perez et al., 2017; Leong et al., 2017) and even whole interactions (e.g., Kinreich et al., 2017). They then compare IBE values between different conditions, or correlate IBE estimates with an outcome variable (e.g., learning) (Leong et al., 2019). For example, Leong and colleagues (2017) compared the amount of observed entrainment across all moments of direct vs averted gaze during 5-minute social interactions: they collapsed all of their data down to a single IBE value per signal frequency band. A similar approach was taken by Perez and colleagues (2017), who compared IBE values estimated separately for different frequency bands and topographical

locations, but again without consideration of how IBE varied over time, and how it may have developed over the course of the interaction.

Effectively, therefore, these approaches produce an index of IBE that includes information on how entrainment varies by frequency (e.g. Leong et al., 2017) and by scalp topography (e.g. Santamaria et al., 2020) – but which excludes information on how IBE fluctuates over time. This omission, we argue, fundamentally hinders our understanding of how real-life infant-adult social interactions are substantiated in the brain. The same observation largely holds for most of the hyperscanning research in adult populations, where similar points have been raised concerning the limitations of current approaches (e.g., Novembre & Ianetti, 2021; Moreau & Dumas, 2021).

2.1.1. The importance of the (missing) temporal dimension

Studies using event-related potentials (ERPs) have shown that even young infants' brains show millisecond-level sensitivity to ostensive signals (e.g., Farroni et al., 2002; Hoehl & Striano, 2008; 2010, Quadrelli et al., 2019). But this research is all unidirectional: it examines how the recipient of an ostensive signal is influenced by the 'sender' of the signal. Very little research has examined the fine-grained temporal dynamics of early social interaction from a bidirectional perspective: by examining how ostensive cues affect the inter-relationship between both partners' brain activity (Wass et al., 2020).

For example, one early study found that, in the 3-9Hz range, neural activity in one partner consistently predicted the other partner's neural activity more strongly during direct compared with indirect gaze (Leong et al., 2017). But how is it mechanistically possible for

two brains to influence each other over such fine-grained temporal scales? To answer this question, it would be useful to know how IBE varies over time *within* periods of direct gaze. This would improve our understanding of how, mechanistically, IBE is established and maintained:

I. First, it is possible that, during social interactions, certain shared behavioural events such as the onsets of periods of direct gaze could drive changes in IBE (see e.g., sections 2.2.2.1 and 2.2.2.2). Here, changes in IBE would result from transient intra brain changes in spectral power (e.g., Grossman et al., 2007) and/or phase (e.g., Rousselet et al., 2007) in both the 'sender' and the 'receiver' of the social cue. For example, this mechanism could be similar to what has been documented for neural entrainment to speech (e.g., Doelling et al., 2014), whereby the onset of the stimulus/ behavioural event drives phasic changes in the brain and leading to increases in entrainment. Beyond speech or vocalisations, mutual gaze onsets, or touch could also act as salient "edges" that create responses in multiple brains at the same time, leading to event-related increases in IBE. According to this model, IBE would be strongly event-locked, peaking immediately after the onset of the behavioural event and decreasing thereafter. The extent of event-locked changes in IBE around behavioural events might also be mediated by other factors including attention (e.g., Golumbic et al., 2013), comprehension (e.g., Pérez et al., 2019), and environmental factors such partner familiarity (e.g., Reindl et al., 2022).

II. Second, it is possible that turn-taking during social interactions could drive changes in IBE. Here, response preparation or anticipation (e.g., Hamilton, 2021; Hirsch et al.,

2017; Kirkland, 2020) and/or mutual sensorimotor predictions (Dikker et al., 2014; Shamay- Tsoory et al., 2019; Hamilton 2021) could lead to concurrent transient changes in either power or phase in both partners (e.g., Mandel et al., 2016; Bögels, 2019), causing changes in IBE that might peak around these 'handover' moments and decrease before and after (e.g., Fig. 2.2d). This mechanism might also be influenced by factors such as the amount (e.g., Nguyen, et al, 2021) and (perceived) quality (e.g., Bloom, 1988) of turn-taking.

III. Third, it is possible that continuous, deterministic intra brain changes, that are not locked to discrete behavioural events but depend on dynamic, gradual changes in the shared environment, lead to gradual, continuous changes in IBE. This mechanism might be driven by intra brain responses to shared cognition and/or mental representations. For example, Simony and colleagues (2016) showed that IBE was increased when participants had a shared understanding of a story (Simony et al., 2016). This mechanism it might also take the form of direct 'neural mimicry'. For example, Kingsbury and colleagues (2019) used *in vivo* electrophysiological recordings to show populations of cells in the dorsomedial PFC that show similar activity when performing an action as when watching it be performed by someone else (Kingsbury et al., 2019). Again, these changes might take the form of changes in power: increases in spectral power throughout an event (e.g., a look/ episode of attention) can increase signal-to-noise ratios and cause changes in sequential IBE (as we show in section 2.3.2.1). This is conceivable as, for example, infant theta power increases through an attentional episode (e.g., Jones et al., 2020). Alternatively, gradual changes in frequency, such as the adjustment of the peak frequency of neural oscillations, could lead to increases in concurrent IBE (section 2.3.2.2). This is

conceivable as, for example, peak alpha frequency can be modulated by task demands (e.g., Samaha & Postle, 2015; Wutz et al., 2018), and recent accounts have theorised that cross-spectrum frequency adjustment at stimulus onset might be a mechanism behind how ERPs are generated (e.g., Burgess 2012). According to this model, entrainment would increase gradually during a social episode. The key differences in model 3 is that this from of IBE might be representative of higher order cognitive functions beyond basic sensorimotor processing

Differentiating between these and other hypotheses is essential to understanding how IBE is achieved and maintained. The aim of this paper is to present algorithms that will allow researchers to address this. In section 2.1.2 we present an overview of key differences between child and adult EEG that are relevant when conducting hyperscanning developmental research. In section 2, we present several measures for estimating concurrent (2.2.2.1, 2.2.2.2) and sequential (2.2.2.3, 2.2.2.4) IBE. Then, in section 3, we illustrate the ability of each metric to capture IBE using simulated data.

2.1.2 Key differences between child and adult EEG

Researchers analysing EEG recorded from infants and children, and EEG recorded using naturalistic paradigms, face several additional challenges as compared to adult EEG researchers that use screen-based paradigms (Noreika et al., 2020).

Firstly, due to increased movements during the recording. This is challenging because signals generated from movements, such as smiling, vocalisations, eye movements, as well as from the neural processing of each of these behaviours, will contribute to the scalp EEG in a

complex way (Georgieva et al., 2020). Although issues of source separation are not new to EEG research, it is known that ICA alone fails to separate different sources in data containing high amounts of movement-related activity (e.g., Plöchl et al., 2012; Dimigen, 2020a). This effect is heightened with infant ICA decompositions, for which it is typically harder to identify which components contain predominately artifactual signals and which contain predominately neural signals, compared to ICA decompositions from adult EEG data (Marriott Haresign et al., 2021 in press). For example, even simple artifacts such as blink artifacts can be more clearly differentiated from the ongoing EEG in adult data, allegedly because these movements are more stereotypical and/or produce artefacts of relatively higher amplitude in adults than in infants (Marriott Haresign et al., 2021). This is even more problematic for naturalistic data, during which eye movements would necessarily be less stereotypical than those produced by participants during screen-based tasks.

Secondly, in a traditional, experimenter-designed paradigm, neural responses are examined relative to experimental events. Although evidence suggests that artifacts in traditional experimenter-designed paradigms are still present, and systematically related to experimental events (Yuval-Greenberg et al., 2008), the fact that the experiment (and so the artifacts) follows a consistent structure means that artifacts are relatively easier to deal with. But in a naturalistic paradigm, the events (e.g., eye gaze onsets) are often less systematically related to the artifacts in the data, as there is no clear and consistent temporal structure between spontaneous events and specific artifacts. The future study of IBE using naturalistic paradigms will need to control for the contributions of non-neural signals in the EEG. It may also treat these non-neural signals as data sources, by looking at entrainment between these movement-related signals, e.g., entrainment between EMG associated with facial affect and vocalisations.

An additional challenge is posed by the intrinsic differences in EEG activity that are observed in recordings from children/ infants compared to adults. For example, we know that the speed at which the brain will process information depends on its maturation (e.g., Taylor et al., 2004) and that the canonical frequency bands in child/infant EEG are typically slower than that of adult EEG. For example, peaks in the power density spectrum associated with alpha activity typically observed in the 8-12Hz range in adults can be seen clearly in one-year-old infant EEG between 6 and 9Hz and are lower still in younger infants (Marshall et al., 2002). This presents a unique problem for developmental researchers interested in phase entrainment between infant and adult EEG. One solution to this problem might be to use cross-frequency entrainment methods (Noreika et al., 2020) for example cross-frequency phase coupling (see section 2.2.2.2.2 for further discussion). It is also known that the amplitude of slower oscillations is larger in infant EEG than in adults', which could arguably affect amplitude-amplitude or amplitude-phase coupling.

2.2. Methods for identifying different types of entrainment between infant and adult EEG data

In this section we present an overview of the different ways of measuring IBE (section 2.2.1) and describe how they can be applied to different aspects of the brain signal (amplitude/ power and phase) (section 2.2.2). In section 2.3 (applied methods) we will guide the reader through application of these methods, using simulated EEG hyperscanning data.

2.2.1 Overview of different ways of measuring inter-brain entrainment (IBE)

Although the term 'entrainment' is sometimes used only to describe only *sequential* relationships between two signals, here we use it in a broader sense, to describe any temporally coordinated relationship between two signals. Inter-brain entrainment (IBE) can, then, be measured in two ways. First, concurrent IBE (see Figure 1) indexes a zero-lag, simultaneous relationship: 'at times when A is high, B is also high' or (for a negative relationship): 'at times when A is high, B is low'. Concurrent IBE is often referred to using the term 'synchrony', and is undirected: A->B is indistinguishable from B->A. Second, sequential IBE indexes a lagged, or temporally oriented relationship: 'changes in A forward-predict changes in B'. Sequential IBE is directed, and as such, unlike concurrent coupling, it can be asymmetrical: it can be true that A forward predicts B without it being true that B forwards-predicts A, and *vice versa*.

IBE can also be measured across both the temporal and the frequency domain, and thus across multiple aspects of the brain signal: amplitude, power, and phase (see Figure 2.1). Power is proportional equal to the amplitude squared and so the two measures are closely related. However, some of the measures we describe (e.g., power correlations, see section 2.2.2.1) can be applied equally to amplitude and/or power, whereas others (e.g., time domain granger causality, see section 2.2.2.3) are applied on amplitude and *not* power and therefore we feel that it is important to distinguish between the two.

Currently, most fNIRS and fMRI hyperscanning studies measure co-fluctuations in the amplitude of the signal – which depending on the method, measures blood oxygenation/deoxygenation (fNIRS), the BOLD signal (fMRI) or voltage (for EEG).

Currently, most EEG hyperscanning studies examine IBE based on the phase of the signal. There appears, however, to be no reason why these trends should not change in the future.

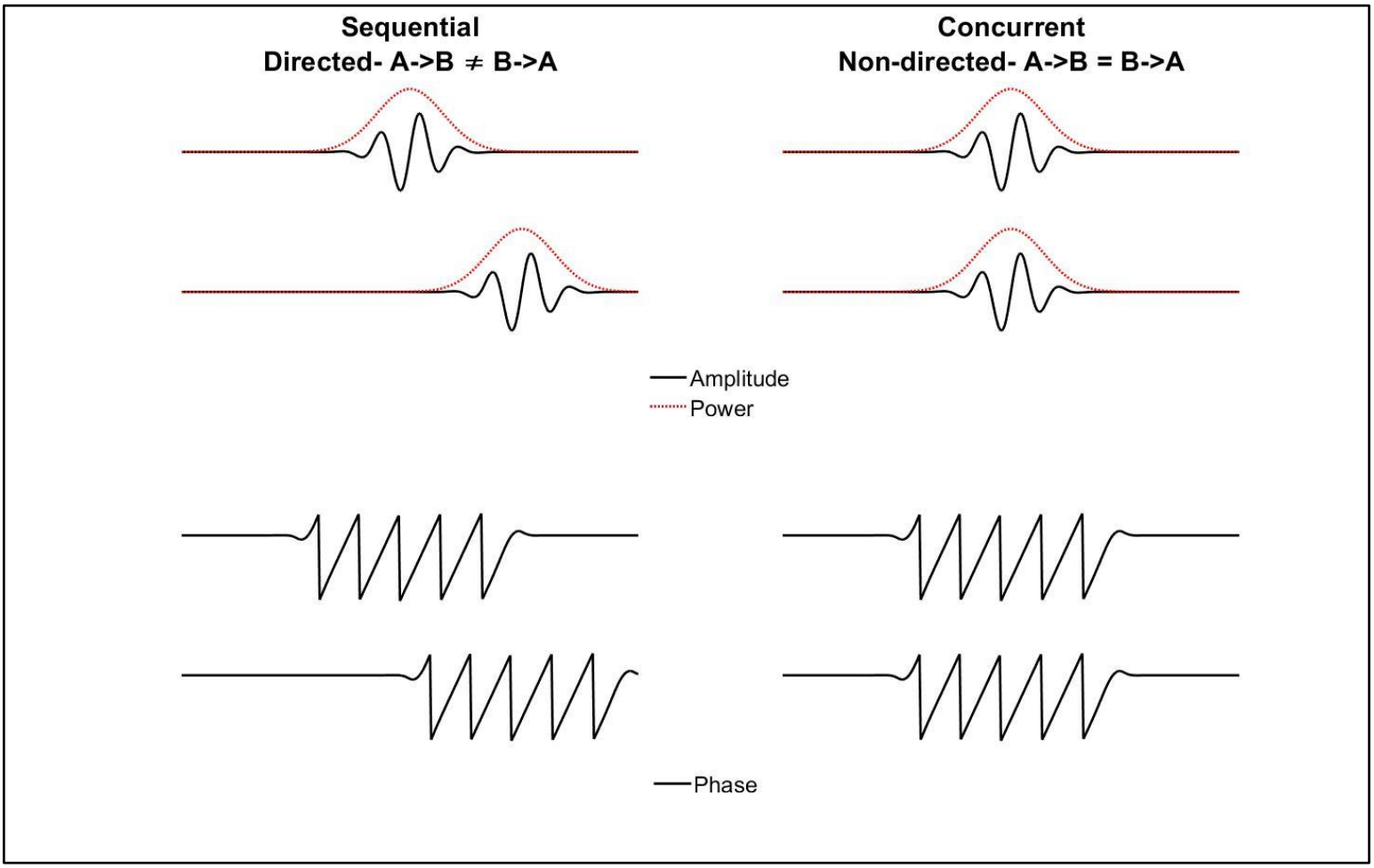


Figure 2.1. Schematic illustration of the two-entrainment metrics, concurrent and sequential, that we consider in the paper, along with the three aspects of the brain signal: amplitude, power and phase.

2.2.2.1 Measuring concurrent IBE of amplitude and power: correlations.

Examining concurrent IBE through correlations is one of the simplest and most flexible techniques. Zero-lag concurrent IBE can simply be measured by calculating correlation coefficients between two time series. Spearman's correlation is generally favoured due to its invariance to non-normally distributed and outlier prone data (Cohen, 2014). The same analysis can be applied either to the amplitude of the brain signal or to the power in particular frequency bands. The accompanying code for this section allows the reader to compute single-trial correlations (Spearman's rho) at each time-frequency point, between pairs of electrodes (e.g., using data from Cz from person one and Cz from person two).

2.2.2.2 Measuring concurrent IBE of phase: ITC and Phase Locking Value (PLV)

Phase locking can be estimated in three ways to detect transient phase changes/shifts (often event-locked), or stable phase-coupling across time. First, point-wise phase consistency (e.g., inter trial coherence, ITC) across repeated events can be estimated over time and electrodes within a single brain (see SM section 8). This produces one estimate of phase consistency per time-point which represents the phase distribution across trials for that time-point and is suitable for detecting transient or discrete event-locked phase changes.

Second point-wise phase locking (e.g., Lachaux et al., 1999) can be estimated according to:

$$PLV_n = \frac{1}{N}\left| \sum_{k=1}^{N} e^{i(\phi(t,k)-\psi(t,k))} \right| \quad (1),$$

Where N is the number of trials, $\phi(t,k)$ is the phase on trial *k*, at time *t*, in channel ϕ and $\psi(t,k)$ at channel ψ. This produces one estimate of phase consistency per time-point which represents the phase locking across trials for that time-point and is suitable for detecting

transient or highly event-locked changes phase locking. PLV_n varies between 0 and 1 where 1 indicates perfect phase locking over trials and 0 indicates no phase locking over trials. Third, phase locking can be also measured within a single trial over a defined temporal window (e.g., Tass et al., 1998). This is useful for estimating whether two oscillations are stably phase-coupling to each other during that window. Phase locking withing a temporal window can be measured according to:

$$PLV_t = \frac{1}{T}\left| \sum_{n=1}^{T} e^{i(\phi(t,n)-\psi(t,n))} \right| \quad (2),$$

Where T is the number of observations or time samples within the window, $\phi(t,n)$ is the phase on observation *n*, at time *t*, in channel ϕ and $\psi(t,n)$ at channel ψ. PLV_t varies between 0 and 1, where 1 indicates perfect phase locking over time and 0 indicates no phase locking over time. Both measures can either be computed between a single brain and an external stimulus with a pseudo-periodic structure (e.g., speech), or estimated at the interpersonal level, between two or more brains. Here the focus is on the latter, though mathematically they are identical. The accompanying code allows for analysis of phase locking within (sliding window) and across trials (for each sample). This will allow researchers to look at how changes in IBE fluctuate over time.

It is worth noting any two signals with a common dominant frequency (e.g., two brains with high power in the alpha band) will show relatively consistent variation in phase over time – and hence high phase locking between the signals. This has been used to argue that PLV is particularly prone to detecting spurious hyper connections (Burgess, 2013). However, this is only a problem in a few scenarios. Real EEG data - even after narrow-band filtering - will

still have random variations in the phase of the signal over time such that PLV between narrowband filtered signals will also change over time. So, two alpha oscillators will show consistently high PLV only where there is little to no random variations in the phase of both signals over time, which is not a very reasonable assumption for real EEG data. In addition, this problem can be at least partially circumvented by relying on permutation techniques. For instance, PLV between two brains can be measured over time-scrambled and real data: if the phase locking is purely attributable to the fact that both brains oscillate at the same rhythm in a rather constant way, the time-scrambled data and the real data will show similar levels of phase-locking. On the contrary, if a substantial part of the phase-locking depends on the real-time interaction between the two partners, real data will show higher phase locking than scrambled data.

2.2.2.2.1 Side note on power and PLV: induced versus evoked responses.

When analysing any event locked changes in EEG power and/or phase-based entrainment it is important to consider whether these are evoked or induced responses. Evoked responses are additive signals superimposed upon the background/ongoing EEG; induced responses are changes in power and/or phase that happen within the background/ongoing EEG. In other words, evoked responses are transient changes that do not relate to background oscillatory activity, while induced responses entirely depend on the adjustment of background oscillators to incoming stimuli. Whilst changes in power/phase resulting from stimulus-locked **evoked** signals could give the appearance of increased entrainment between two brains, this is interpretationally quite different to potential changes due to **induced** neural activity driving increases in entrainment. For example, if increases in spectral power from two signals are driven purely by evoked and not induced responses then it is incorrect to examine phase resetting as a potential mechanism behind IBE (or phase-locking more

generally) and incorrect to use the term neural IBE to refer to these mechanisms (e.g., Keitel et al., 2021).

This problem is further complicated because, as Muthukumaraswamy and colleagues (2011) show, transient increases in power can lower error in phase estimation and give the appearance of heightened phase locking (see also, Burgess, 2013). Separating increases in power from genuine increases in phase locking is difficult and continually debated (e.g., Sauseng et al., 2007). As a consequence, the best practice for researchers using event-related phase-locking is to always show accompanying power plots and examine power and phase simultaneously.

2.2.2.2.2 Side note on cross-frequency PLV: dealing with different canonical frequencies.

As described above (section 1.2) the canonical frequency bands in infant EEG are typically slower compared to adult EEG. It may, therefore, be more appropriate for researchers measuring the quantity of phase-locking between infant and adult EEG to use cross-frequency phase locking. Cross frequency phase entrainment or PLV *m:n* is calculated similarly to PLV as follows:

$$PLV_{mn} = \frac{1}{N}\left| \sum_{k=1}^{N} e^{i(\Delta\phi_k(f_n, f_m, t, k))} \right| \quad (3),$$

Where, *N* is the number of trials and $\Delta\phi_k(f_n, f_m, t, k)$ is calculated as follows:

$$\Delta\phi_k(f_n, f_m, t) = \left(\frac{n+m}{2 \cdot m} \cdot \phi(f_m, t, k) - \frac{m+n}{2 \cdot n} \cdot \psi(f_n, t, k)\right) \quad (4),$$

Where *n* and *m* are the centre frequencies of the two signals and should be integer values satisfying the equation $m \cdot f_n = n \cdot f_m$, and $\phi(f_m, t)$, is the phase angle at channel ϕ, at time *t,* on trial *k*, and channel ψ. Note that as we have described in section 2.2.2.2 PLV can be applied over trials (1) or in a time window within a given trial (2). The same applies for cross frequency PLV, although here we only describe the equation for estimating cross frequency PLV over trials (3). Cross frequency phase locking shares the same underlying interpretation as standard phase locking. In accompanying articles in the special edition (Kayhan et al., 2022), we have provided readers with a full pipeline for computing cross-frequency phase-locking within (sliding window) and across trials (for each sample).

2.2.2.3 Measuring sequential IBE of amplitude and power – Granger Causality (GC)

The simplest way to measure sequential IBE is simply to repeat the Spearman's correlation described in section 2.2.2.1 while shifting one time series forwards or backwards in time relative to the other. For example, if we find that the correlation between two time series x and y is stronger when time series x is backward shifted with respect to time series y, compared with when the simultaneous ('zero-lag') correlation between the two-time series is examined, then this indicates that changes in x tend to predict changes in y.

Granger Causality is closely related to this approach, but as well as looking at the time-lagged relationship between two time series, it also increases the sensitivity of the prediction by considering how one time-series forwards predict itself over time (known as the autocorrelation). Given two-time series x and y, Granger Causality is a measure of the extent to which time series x can be predicted by previous samples of y above and beyond how well time series x can be predicted by previous samples of x alone.

GC is defined through the log of ratios of error terms between the bivariate and univariate regressive models, following:

$$GC = \ln\left(\frac{\mathrm{var}(e_x)}{\mathrm{var}(e_{xy})}\right) \quad (5),$$

Where e_x is the error term obtained from the univariate autoregressive model fit and e_{xy} is the error term obtained from the bivariate regressive model fit. Again, the same approach can be adopted to look either at the amplitude of the brain signal, or at the power within frequency bands. Time frequency (spectral) GC involves computing the dot product between the regressive coefficients and complex sine waves, (analogous to the Fourier transform) and then applying those results to the error variance via the transfer function (Cohen, 2014).

Finally, Partially Directed Coherence (PDC) is a frequency domain formulation of GC (Sameshima and Baccala, 2014), measured from the coefficients derived from the autoregressive modelling process described above. PDC has also been used to investigate IBE across adult - infant dyads (e.g., Leong et al., 2017; Santamaria et al., 2020). PDC along with other methods of frequency domain entrainment based on autoregressive modelling can be implemented using the extended multivariate autoregressive modelling toolbox (Faes, et al., 2013).

2.2.2.3.1 Side note on EEG data stationarity and GC

Stationarity refers to whether the statistical properties of data change over time. For example, EEG data that contain low-frequency drifts over time can cause data to become nonstationary as the mean of the data is changing over time. Non-stationarity may take various forms. One

form of non-stationarity manifest itself through unit root processes - where the data may exhibit a stochastic trend or “random drift”. This form of non-stationarity is common in financial time series (e.g., stock prices), but less common in neuroimaging data (because neurophysiological processes are generally physically constrained). Many stationarity tests (including the KPSS test which is implemented within the MVGC toolbox (Barnett and Seth, 2014)) only test for unit-root stationarity. However, unit-root is not the only kind of non-stationarity; there is also "structurally varying" non-stationarity, which may be more common in ERP or more generally task-related data. Here, the statistical properties of the time series, e.g., mean/standard deviation over time, the amount of periodicity in the data, change over time, either in a structured/deterministic or stochastic way (unit root tests may or may not fail on this form of non-stationarity). This implies that the Granger causality itself may change over time. Overall stationarity in neuroscience for GC is an ongoing problem (e.g., see Barnett et al., 2018a; 2018b). Current common approaches to address non-stationarity in EEG data include polynomial detrending (Seth et al., 2015), and/or subtraction of the averaged ERP from single-trial data (e.g., Wang et al., 2008), but these are limited. Another viable solution to address nonstationary EEG data is to segment the data into shorter time windows in which the data would be stationary enough to perform GC analysis, although this approach needs more empirical testing.

2.2.2.3.2 Side note on model order and GC

A crucial parameter to consider when using GC analysis is model order. Model order determines the number of previous samples of a time series that will be used in the (bivariate) autoregressive model fit. For instance, if your data is sampled at 1000Hz, a model order of 5 means that the model will use a weighted sum of the previous 5ms of data. The model order used will in part determine the frequency precision of the spectral GC estimates. In our

example using only 5ms of data, we would only base our GC estimate on 1/50th of a cycle of 4Hz activity. To better capture low-frequency dynamics, it can therefore often be useful to downsample the data prior to analysis. For example, if we resampled our data from 512Hz to 128Hz, still using a model order of 5 would mean we now consider the previous 39ms of data in the model fit. Alternatively, one could increase the model order, but models with higher orders typically require longer time segments and more trials as there are more parameters that need to be estimated (Cohen, 2014). Further considering lower frequency dynamics is a good reason to lean towards the highest model order that is appropriate for the data. In our example increasing the model order to 15 would result in us considering the previous 120ms (at 128Hz), almost half a cycle of 4Hz activity. This example illustrates the fact that the model order used can have important consequences on the estimation of GC. Some routines exist to help guide the estimation of model order, the most common being Bayes information criterion (BIC) and Akiake information criterion (AIC). Both are implemented within the MVGC toolbox (Barnett and Seth, 2014).

2.2.2.3.3 Side note on spectral power and GC

The relationship between spectral power and spectral GC is still uncertain, for example at present it is unclear how changes in spectral power affect GC estimates and anecdotal evidence suggest that it is not uncommon to find correlations between spectral power and (spectral) Granger Causality. Empirical research has shown that increases in event-locked spectral power corresponding to ERPs co-occur with increases in spectral GC (e.g., Wang et al., 2008), but whether these changes in power *caused* the changes in GC or vice versa is uncertain. It will be important for future research to fully explore this relationship (e.g., Winkler et al., 2015). For example, does the strength of the GC scales linearly with the

amount of spectral power? And how is this relationship affected by the sampling rate, signal to noise ratios and so on?

2.2.2.4 Measuring sequential IBE of phase – Phase Transfer Entropy (PTE)

Phase transfer entropy (PTE) allows researchers to measure sequential IBE of phase. It is calculated using the following equation:

$$\text{Phase } TE_{x\to y} = H(\theta_y(t), \theta_y(t')) + H(\theta_y(t'), \theta_x(t')) - H(\theta_y(t')) - H(\theta_y(t), \theta_y(t'), \theta_x(t')) \quad (6),$$

Where $\theta(t)$ is the phase of signal X(t), t' = t – δ, and $\theta_x(t')$ and $\theta_y(t')$ are the previous states of the phase angle time series of x and y, with a given lag of δ. Given two-time series x and y, like GC, transfer entropy (TE) estimates whether including the past of x influences our ability to predict y and vice versa. However, unlike GC, TE does this by comparing conditional probabilities (Lobier et al., 2014). E.g., if a signal X 'causes'/ 'disambiguates' a signal Y, then the probability density of the future of Y conditioned on its past should be different from the probability density of the future of Y conditioned on the pasts of both X and Y. As transfer entropy is based on the same underlying principles as GC, it has been shown that results obtained using GC and PTE are identical for Gaussian variables (e.g., Barnett et al., 2009; Barnett and Bossomaier, 2013). Therefore, results from phase transfer entropy analyses may be interpreted as reflecting directed information flow between two phase angle time series.

Whilst Phase Transfer Entropy has not been widely used within cognitive neuroscience as a framework for analysing entrainment patterns between two systems, it has many advantages

and useful properties. For example, as entropy is not based on the temporal structure of the data, it can be computed over time *and* trials, whereas other measures such as PLV can only be computed over time *or* trials, not both. This is a major advantage as including data from time and trials simultaneously means that entropy can be computed in shorter time windows than other window-based entrainment measures (e.g., GC), thus retaining a greater degree of the original temporal precision of the data whilst still having sufficiently high signal to noise ratios (Cohen, 2014).

2.2.3. Cautionary note on the importance of the temporal scale.

Many of the metrics described above are highly sensitive to the temporal scale of the analysis. For example, if we observed a transient increase in spectral power in two signals (x and y), where the peak of y occurred a few hundred milliseconds after the peak of x. When using a fine temporal scale (e.g., estimating entrainment in a 200ms sliding window), we would not detect changes in concurrent IBE, but if we were to use a larger time window (e.g., estimating entrainment in a 1s sliding window) it is possible to observe changes in concurrent IBE (e.g., see figure 6). This we illustrate using simulated data in section 2.3.3. This is because, although downsampling can be a useful step in some analysis (e.g., for GC, see section 2.2.2.3.2), downsampling can create artifacts by spreading the signal in time that can lead to the detection of spurious entrainment.

2.3. Simulations and applied methods for measuring IBE and differentiating event-locked from non-event-locked changes.

In this section we present simulations in which we artificially introduced a given relationship between two time series. We then compute the metrics described above to assess how well each metric reflects this relationship. The purpose of this is to guide the reader through application of the various metrics. Throughout the section, we also discuss how different signal changes can manifest either as transient, event-locked changes (section 2.3.1) or as more continuous, non-event-locked changes (section 2.3.2). In sections 2.3.4 and 2.3.5 we describe methods to quantify whether observed event-locked changes in IBE differ significantly from chance.

2.3.1 Simulations - event-locked changes in child-adult neural entrainment

2.3.1.1 Amplitude and power

At its most simple level, computing the correlation coefficients for time-frequency amplitude/power does not involve any algorithms more complex than a Spearman's correlation. In the accompanying code, we provide routines for computing concurrent amplitude/power entrainment. However, for the remainder of this section we will focus on GC, which is a more appropriate complex measure for assessing entrainment in EEG data: as detailed above, it not only assesses the potential contribution of one signal to another over time (e.g., the extent to which time series x can be predicted by previous samples of y), but also considers autocorrelations between each signal (e.g., the extent to which x can be predicted by previous samples of x alone).

To illustrate how event-locked neural responses might give rise to changes in sequential amplitude/power IBE, we simulated two ERP-like signals (x and y) (see Figure 2.2). Signal y

was generated from previous samples of x plus noise (see SM sections 2.6.2 and 2.6.7 for full details). From this simulation, it can be seen that the sequential IBE between x and y observable in the raw data (Fig 2.2a) manifests as strong x->y GC influences but not y->x GC influences, as expected (Fig 2.2d). When the same analysis is applied to the power of the signal (Fig 2.2b, 2.2c), the predicted results are again observed. Spectral x->y GC influences are observed across a range of lower frequencies (Fig 2.2e), but no spectral y->x effects are observed (Fig 2.2f). In the accompanying code for this section, we provide the user with routines for implementation of time domain and spectral GC (i.e., sequential IBE based on amplitude/power) for measuring event locked changes in EEG hyperscanning data. The user is also able to easily specify more advanced parameters such as the time window size and model order used for the time-varying GC estimates.

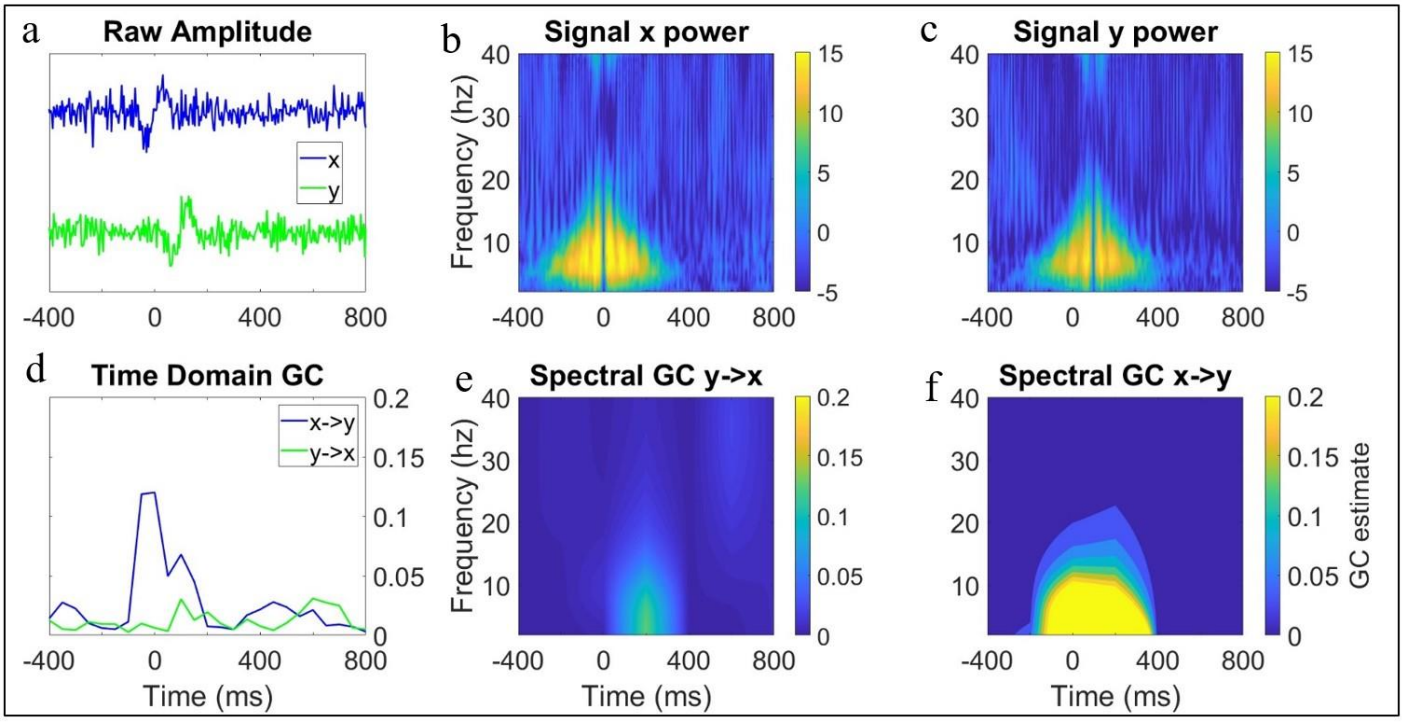


Figure 2.2. Simulated data showing one mechanism that could give rise to increases in interbrain granger causality between parents and infants. (a) shows the two correlated (single-trial amplitude) transient signals x and y. Y was generated from previous samples of x with a lag of 100ms, such that (d) there is a substantial event locked increase in GC from x to y but no GC influence from y to x. (b) shows time-frequency power from signal x from panel a. (c) shows time-frequency power from signal y from panel a. (e) shows spectral GC from y to x and (g) from x to y.

2.3.1.2 Phase

To illustrate how event-locked neural responses might give rise to changes in concurrent phase-based IBE, we simulated 100 trials of two partially phase-locked signals (x and y) with a concurrent phase reset/modulation +200ms after an event (time 0) (see Fig 2.3a) (see SM section 4 for more details). From this simulation, it can be seen that, during the time window following the manipulation at +200ms, the phase angles of the two time series converge (Fig 2.3a) as expected. The phase locking values of the two time series also converge (Fig 2.3c) as expected.

To illustrate the sequential phase IBE, we simulated 100 trials of x and y again in the same way, but here the phase modulation in y occurred 200ms later than in signal x (see Fig 2.3b) (see SM section 2.6.5 for more details). From this simulation it can be seen that when phase modulations in one signal occur later/ earlier than phase modulations in another signal (e.g., x in Fig 2.3b becomes phase-locked at +200ms, and y in Fig 2.3b becomes phase-locked at +400ms) and that these modulations are correlated, then this relationship (or form of sequential IBE) can be captured using directed phase IBE methods such as phase transfer

entropy – illustrated by the increase in PTE in the time window between phase modulations of x and y (~+200-400ms) (Fig 2.3d). Also note from the simulation that as y becomes phase locked (or entrains to x) at +400ms this causes an increase in PTE from y to x. This is because both signals were generated from pure sine waves plus noise and so when y also becomes phase locked at +400ms, the activity in y is also now predictive of the activity in x.

In the accompanying code for this section, we provide the user with routines for full implementations of inter-individual, time-frequency PLV and PTE (i.e., sequential IBE based on phase). The user can easily specify parameters such as time window size for PLV/PTE as well as the model order for PTE.

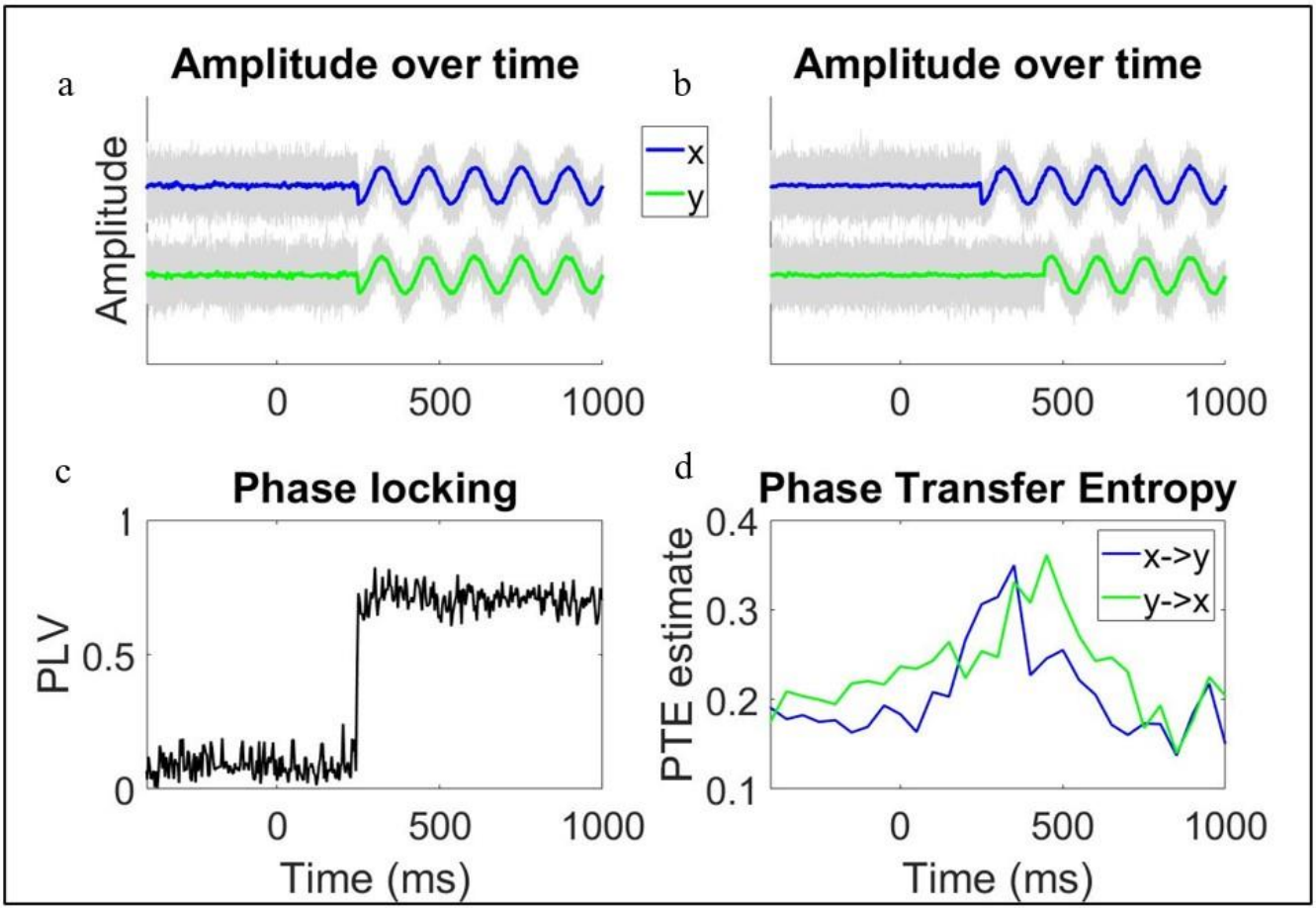


Figure 2.3. Simulated data showing how event locked phase modulations could give rise to phase-based IBE between two brains. (a) Time series data x and y were subjected to a phase reset at +200ms so that they become purely phase locked. Note grey lines show single trial data while blue and green lines indicate data averaged over trials. The increased consistency in phase angle circa +200ms between x and y yields (c) a notable increase in the phase-locking between x and y ~200ms. (b) Simulated data showing a situation in which the phase modulations in one signal (x) predict the phase modulations in another signal (y) and how this lagged/ directed relationship in phase can be captured (d) using phase transfer entropy.

2.3.2 Simulations – non-event-locked changes

2.3.2.1 Amplitude and power

To illustrate how gradual changes in amplitude/power IBE that are not time-locked to the onset of an event might arise, we simulated two oscillatory signals (x and y) where y was generated from previous samples of x plus white noise (see Fig 2.4a). To simulate a gradual change in GC we reduced this noise parameter over time (see SM section 3 and 7 for more details). From this simulation, it can be seen that, as expected, the x->y GC influence increases during the time window, but no changes in y->x GC influences are observed (Fig 2.4b).

In the accompanying code for this section, we show how the same code from the previous section (2.3.1.1) can be leveraged to look at questions regarding non-event locked changes in inter brain IBE. No new algorithms are implemented here.

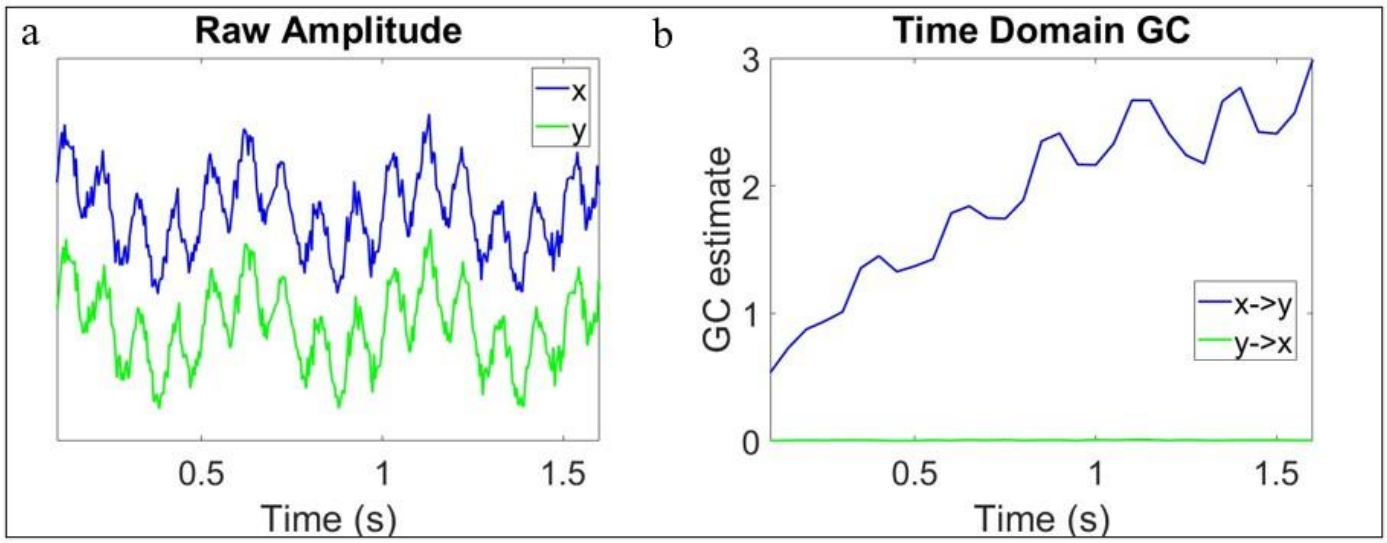


Figure 2.4. Simulated data showing one mechanism that that could give rise to increases in interbrain granger causality between parents and infants. (a) shows two oscillatory signals. Y was generated as a product of previous samples of x with a lag of 25ms. We decreased the amount of noise in x over time to simulate (b) a gradual increase in GC from x to y throughout the segment. Details given in supporting materials for chapter 2, appendix A, section 2.6.3.

2.3.2.2 Phase

To illustrate how gradual changes in phase IBE might arise that are not time-locked to the onset of an event, we simulated two oscillatory signals (x and y) with slow drifts in peak frequency over time (signal x linearly increased in peak frequency from 6 to 9Hz and signal y

decreased from 12-9Hz) (see Figure 2.5). Full details of how we simulated this data, and the time-frequency decomposition can be found in the supplementary materials (SM 2.6.6 and 2.6.7). From this simulation, it can be seen that the closer the signals become in peak frequency the more consistent the relationship between phase angles over time is, and thus the higher the phase-locking value between x and y is.

In the accompanying code for this section, we show how the same code from the previous section (2.3.1.2) can be leveraged to look at questions regarding non-event locked changes in inter brain entrainment. No new algorithms are implemented here.

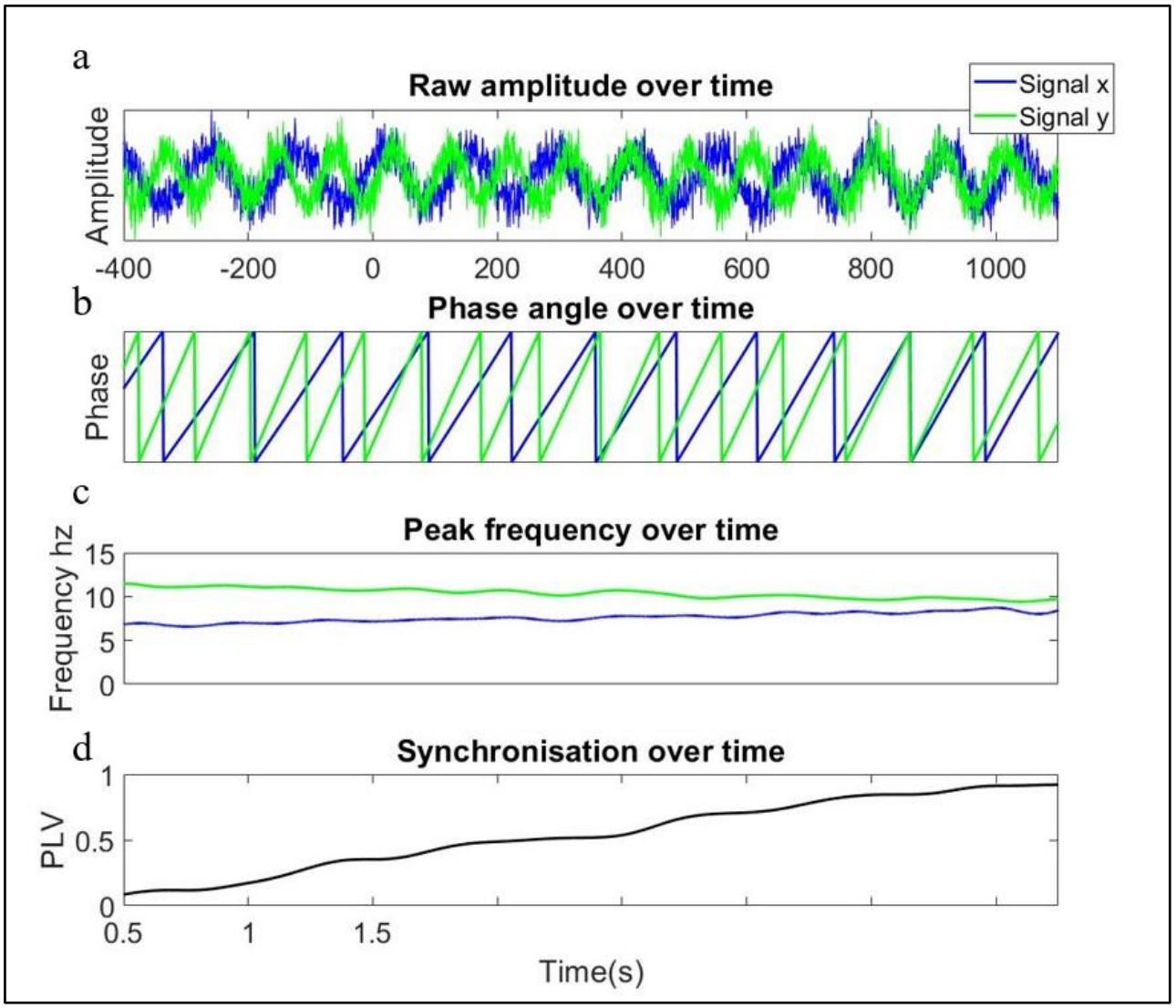


Figure 2.5. Simulated data showing one mechanism that could give rise to increases interbrain phase locking between parents and infants. (a) Time series data x and y both exhibit slow trends in frequency over time toward a common peak frequency. (c) signal x increases from 6 to 9Hz over the time course whereas time series y decreases from 12 to 9Hz. The closer the signals become in peak frequency (b) the more consistent the relationship between phase angles over time is, (d) yielding a gradual increase in PLV between x and y over time.

2.3.3. Simulation- cautionary note on the importance of the temporal scale.

In this section we highlight how the temporal scale of the analysis influences concurrent IBE estimates. To illustrate this, we simulated two signals that show an event-locked transient increase in spectral power which peaks 300ms later in signal y compared with signal x (Figure 2.6a, 2.6b). Details of the time-frequency decomposition can be found in SM 7. To compute concurrent IBE, we performed two calculations: first, we calculated Spearman's correlations between the power of time series x and y at each time-frequency point independently (see Fig 2.6e). Second, we down sampled the data using a 0.5s sliding window with 200ms of overlap between successive windows (Figs 2.6c, 2.6d) before repeating the same analysis (Fig 2.6f). When using a fine temporal scale, we detect no changes in concurrent IBE, but when using a larger time window (reduced temporal precision) we do (Fig 2.6f). This illustrates how the pre-processing of data prior to IBE analyses can alter the results.

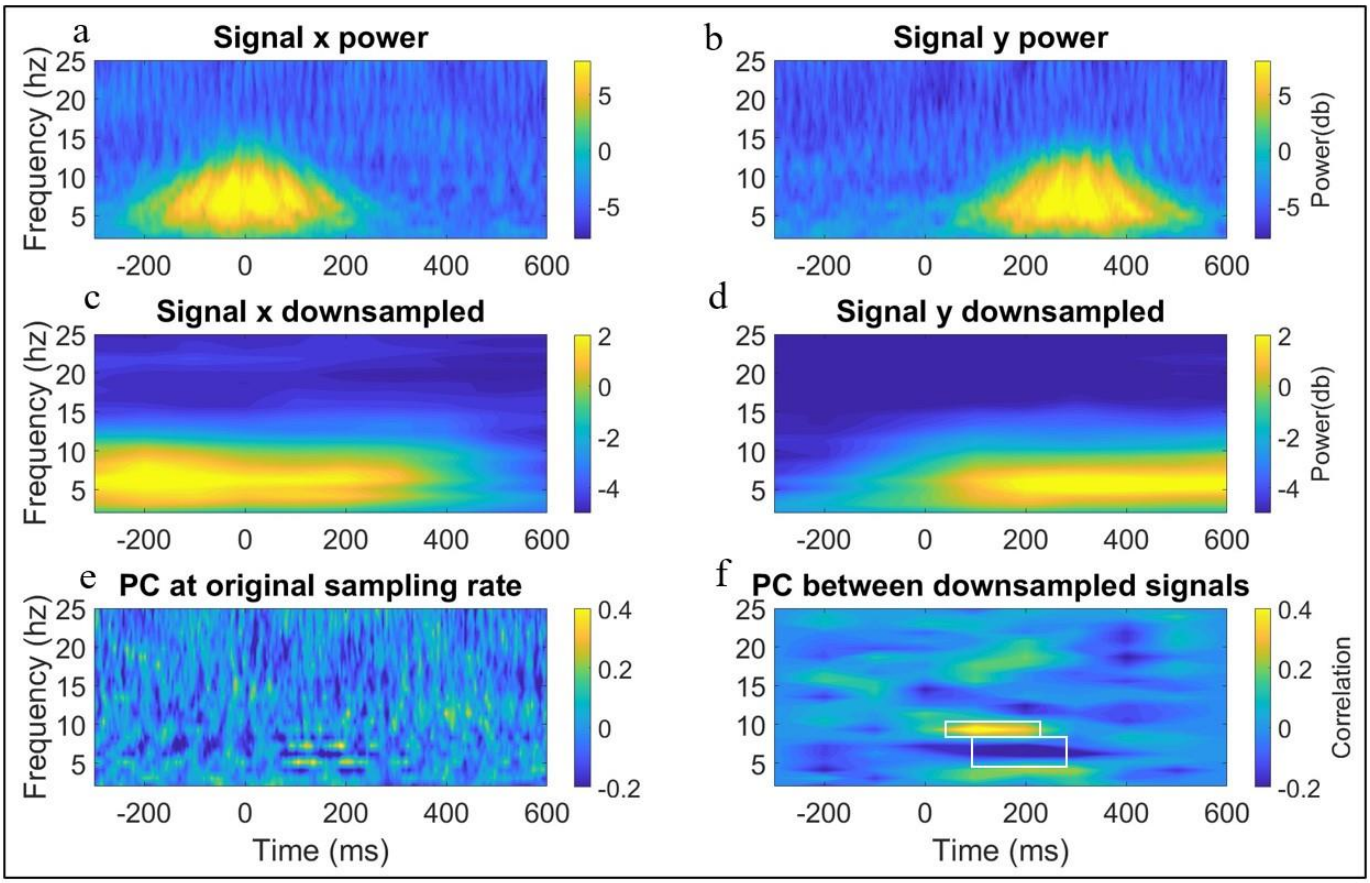


Figure 2.6. Simulation illustrating the importance of using an appropriate time window for calculating concurrent entrainment. Panel a shows the time-frequency power of signal x. Panel b shows the time-frequency power of signal y. Panel c shows the down sampled (using moving window average) time-frequency power of signal x. Panel d shows the down sampled, time-frequency power of signal x. Panel e shows concurrent IBE (spearman's correlation of single trial power (PC) between x and y) computed at each time-frequency point (i.e., original temporal scale of data). Panel f shows the same concurrent IBE but computed on the down sampled data. The AOIs on panel f indicate regions of significant correlations. This figure shows that down sampling the data using sliding window averages can change the observed temporal dynamics of inter-brain entrainment.

2.3.4 Quantifying event-locked changes in child adult neural entrainment

In this final section we consider the question of statistical significance. If we know that specific events occurred in the data, how can we test whether statistically significant changes in IBE occurred relative to these events?

2.3.4.1 Amplitude and power

Significant changes in GC can be evaluated with F-statistics, which is implemented through the MVGC toolbox (Barnett & Seth, 2014). Statistical significance can also be obtained via nonparametric permutation testing, which can be applied to power correlations, time domain and spectral GC (Maris & Oostenveld, 2007). The benefit of the latter measure is that it also deals with the problem of multiple comparisons. For measuring changes in GC that are strongly time/event locked, permuting the order of the time segments within trials is generally recommended over permuting the trial order whilst leaving the time segments intact (Cohen, 2014).

2.3.4.2 Phase

Assuming a von Mises distribution (normal distribution for circular data) statistical significance of ITC and PLV can be evaluated against a p-value, approximated against the null hypothesis using the Rayleigh's test which can be implemented using the Circstat toolbox (Berens, 2009). Statistical significance can also be assessed against a threshold ITC/PLV value (Cohen, 2014). Any values which exceed this resulting threshold can be considered significant. Alternatively, the significance of time-frequency varying ITC and

PLV as well as PTE (when computed in a sliding window within trials) can also be assessed using nonparametric permutation testing (e.g., Maris & Oostenveld, 2007).

2.3.5. Correcting for multiple comparisons when measuring changes in parent-child neural entrainment.

When analysing EEG data, we are typically interested in how a given effect varies as a product of time, frequency, and topography. This makes exploratory EEG analysis susceptible to the problem of multiple comparisons: increasing the number of statistical inferences drawn from the data will also increase the likelihood of obtaining a significant result. There are several approaches to correct for this problem. For example, if you are only testing a limited number of regions/frequencies of interest it is appropriate to use the Bonferroni correction method. For more complex comparisons involving a large number of channels and time-frequency points, Bonferroni correction is not appropriate. In these situations, correction for multiple comparisons should be made using pixel or cluster-based permutation statistics (Maris & Oostenveld, 2007). For a more detailed discussion of which correction methods to use when we refer the author to Cohen (2014, chapter 33).

2.4. Discussion

Procambarus clarkia, a breed of freshwater crayfish, exhibit only a small range of social behaviours, primarily focussed on dominance/ subordination, yet their physiological systems are capable of supporting these interactions as well as intra-individual interactions with their environment with a remarkable level of temporal fidelity (e.g., Schapker et al., 2002). The way that humans interact with their environment and each other is infinitely more complex

and multi-layered (Hasson & Frith, 2016; Hoehl et al., 2021; Murray et al., 2016). However, most researchers who study interacting humans during social engagement typically do so using methods that measure how entrainment varies by frequency/topography and between different experimental conditions (or participants), but which obscure how entrainment varies over time. In this article, we have argued that this omission prevents us from developing a mechanistic understanding of how neural entrainment is established and maintained.

In this article, we presented algorithms that allow researchers to measure how entrainment between two brains (or more generally physiological systems) varies as a function of time. We have differentiated between two types of entrainment (section 2.2.2): concurrent ('when A is high, B is high') and sequential entrainment ('changes in A forward-predict changes in B'). And we have described how these measures can be applied to three aspects of the neural signal: amplitude, power and phase (section 2.2.2, see Figure 2.2.1).

We hope that this guided simulation study and tutorial will help facilitate further research into the possible mechanisms underpinning child-adult neural entrainment. For example, measuring changes in concurrent entrainment of amplitude and power using correlations (section 2.2.2.1), or of phase using PLV (section 2.2.2.2) might be used to explore the possibility that certain behavioural events during social interactions could lead to local increases in IBE, for example around moments of mutual gaze onsets or vocalisations (see section 2.1.1).

Further measuring changes in sequential entrainment of amplitude and power using GC (section 2.2.2.3), or of phase using PLV (section 2.2.2.2) or PTE (section 2.2.2.4), might be used to explore the possibility that response preparation or anticipation and mutual prediction

might lead to changes in sequential IBE (see section 2.1.1). This could in theory involve concurrent, transient changes in either power or phase in both partners (e.g., Mandel et al., 2016; Bögels, 2019), that would drive changes in sequential IBE.

Lastly, the same methods described above might be used to explore the possibility that continuous intra brain changes, that are not locked to behavioural events, could also lead to gradual changes in IBE. This might be substantiated for example, through shared cognition and/or mental representations or through direct 'neural mimicry', even in the absence of explicit turn-taking (Hamilton, 2021; Kingsbury et al., 2019). These concepts are discussed further in section 2.1.1. Alternatively, gradual changes in phase, such as the adjustment of the peak frequency of neural oscillations, could lead to increases in concurrent IBE of phase (section 2.3.2.2), through, for example, concomitant modulations of peak alpha frequency in response to task demands (e.g., Samaha & Postle, 2015; Wutz et al., 2018).

Overall, the study of child-adult neural entrainment is still in its infancy and many very basic questions regarding how changes in inter brain entrainment are substantiated at the neural level, and are mediated through behaviour, remain unanswered. It is our hope that the material presented in this paper will aid researchers in addressing these fundamental questions.

2.4.1. Outstanding issues

There are, of course, many outstanding issues with the analysis of EEG hyperscanning data. Substantial questions remain about how successfully artifacts can be removed from brain data (section 2.1); in understanding the relationship between power changes and IBE (sections

2.2.2.2.1 and 2.2.2.3.3) and the problems associated with non-stationarity in EEG data and Granger Causal analyses (section 2.2.2.3.1); and so on.

Two further outstanding issues should be noted. First, we are often considering events that have different periodic structures, and that unfold over different time scales. For example, researchers might want to examine the relationship between eye gaze shifts (which take place every ~300ms – i.e., at ~3Hz), changes in autonomic arousal (between ~0.01and ~0.5Hz) and changes in EEG (between ~2Hz-~30Hz). Although we have presented some methods for looking at this – such as temporal correlations in power at different frequencies (section 2.2.2.1) and cross-frequency PLV (section 2.2.2.2.2) – we have not discussed other approaches, such as phase-amplitude coupling (Canolty & Knight, 2010; Tort et al., 2010) that would also be useful.

Second, we have concentrated exclusivity on IBE in relation to bivariate behaviours (e.g., mutual gaze). Unlike the ways in which social information processing is typically studied (i.e., using repeated, discrete and unecological screen-based stimulus), real social interactions involve highly layered and complex sequences of multimodal events that can unfold over multiple time scales in a continuous and interdependent way. For example, consider the multimodal pathways to joint attention as illustrated by Yu and Smith (2016), in which sequences of social interactions between parents and infants often involve initiating and responding to various postural and gestural movements, as well as visual (gaze) information and vocalisations, presented in combination. Future work will require more advanced data analysis and the collection of larger datasets, to explore IBE in relation to more complex multivariate behavioural datasets (for example those modelling gaze, touch, affect and vocal data simultaneously).

2.5. Conclusion

The focus on fine-grained neural responses that we advocated in this paper has the potential to provide valuable new insights into the neural processes that support dynamic social interaction, beyond what is possible using the current/ standard approaches adopted in EEG hyperscanning studies. It is our hope that the considerations that we highlighted, and the methods that we described, will pave the way for future studies which will analyse in more depth the rich temporal dynamics of neural activity, bringing us closer to the true complexity of brain functioning.

2.6. Supporting materials for chapter 2

Appendix A

2.6.1. Simulation for event locked changes in concurrent amplitude/ power entrainment

To illustrate how event locked neural responses might give rise to changes in concurrent amplitude/power inter brain entrainment (IBE) (see section 2.3.1.1 of main text), we simulated two time series ‘x’ and ‘y’ each embedded with one purely phase-locked transient oscillation as a product of a sine wave (640 samples, at a sampling rate of 256 Hz) at 7Hz convolved with a gaussian kernel (640 samples, sd of 0.1). The peak of the transient oscillation of time series y was offset from the peak of time series x by +100ms. This was done to be practically most similar to the examples on measuring changes in sequential amplitude/ power-based entrainment (e.g., see figure 2 of main text), but to still highlight how changes in event locked concurrent power entrainment can be measured at a fine temporal scale (e.g., looking at sub-second changes). We simulated 100 trials of data in this way. Details on how we derived single trial EEG power can be found in section 7. Spearman's correlation was calculated between signals x and y at each time-frequency point.

2.6.2. Simulation for event locked changes in sequential amplitude/ power entrainment

To illustrate how event-locked neural responses might give rise to changes in sequential amplitude/power IBE (see section 2.3.1.1. of main text), we simulated two time series x and y each embedded with one transient oscillation as a product of a sine wave (640 time points, at a sampling rate of 256 Hz) at 7Hz convolved with a gaussian kernel (640 samples, sd of 0.1). Time-series y was generated as a product of weighted previous samples of x, according to:

$$y_t = \psi x_{t-\mathrm{n}} \qquad (1),$$

where y_t is the current value of y, ψ is an autoregressive coefficient (set at 1 in this example) and $x_{t-\mathrm{n}}$ is the previous sample of x, at a given lag n. To keep the model simplistic, a model order of 1 was used, meaning that time series y was generated using one previous sample of time series x. Both signals were then mixed with white noise with a standard deviation of 0.5. We simulated 100 trials in this way. In each trial, the amplitude of the sine wave used to generate the transient oscillation in x and y was set at 3 times a random number drawn from a normal distribution. This was to simulate the data being more physiologically 'realistic' as well as yielding an increase in time domain and spectral GC, due to the single trial amplitudes being highly correlated. Single trail time-frequency power was computed in the same way as in the simulation of concurrent power entrainment (see above). To compute GC, we used functionality from the MVGC toolbox (Barnett and Seth, 2014). In the accompanying code, we provide routines on implementing code from the toolbox with EEG hyperscanning data, to compute GC estimates for amplitude and power. As we generated the peak of signal y to be forward lagged by +100ms relative to the peak of x we can expect to

see a relative increase in the time domain and spectral GC in a 100ms time window between peaks (in time) of signals x and y. We also expect that, as both x and y are transient signals and both signals return to baseline after 100ms, we will see no other significant changes in GC. Further, as x was not defined by y, we expect a unidirectional influence only from x to y. The temporal precision of the GC results is determined by the number of samples of the window used to compute GC. Previous literature has used window lengths between 50ms (Ding Bressler et al., 2000) and 2s (Barret et al., 2012), although the most appropriate window length will depend on individual task-related design. In this example, GC was computed in a 200ms sliding window, moved in increments of 50ms. The model order used in the autoregressive model fit for the GC estimates was 1.

2.6.3. Simulation for non-event locked changes in sequential amplitude/ power entrainment

To illustrate how non-event-locked neural responses might give rise to changes in sequential amplitude/power IBE (see section 2.3.2.1 of main text) we simulated one time series as a mixture of one 'slow' sine wave at 2Hz and one 'fast' sine wave at 10Hz (640 time points, at a sampling rate of 256 Hz for both) and white noise (sd 0.5)- labelled as time series x. Time series y was generated as a product of weighted previous samples of x plus white noise with a standard deviation of 0.5, following the same procedure as in the previous simulation. In this example, a model order of 1 was used. We simulated 100 trials in this way**.** As y was generated from previous samples of x plus noise (random variation), to simulate a gradual increase in GC over time we linearly reduced the magnitude of noise parameter/ random variation in the system, such that the values of y became closer to the 'true'/ real signal elements of x, without the confounding noise. GC was computed in the same way and using

the same code as provided for the above simulation. The model order used in the autoregressive model fit for the GC estimates was 1.

2.6.4. Simulation for event locked changes in concurrent phase entrainment

To illustrate how event-locked neural responses might give rise to changes in concurrent phase-based IBE (see section 2.3.1.2 of main text), we simulated two-time series (640 samples at 256 Hz) as partially phase-locked signals (e.g., non-phase locked before time +200ms and purely phase locked after time = 200ms). At time =+200ms we simulated a phase-modulated such that at subsequent time points the signals were purely phase-locked. We simulated 100 trials of data in this way. To calculate the frequency-specific phase-locking value used in our examples (see section 2.3.1.2 of main text) both signals were filtered between 6 and 9 Hz using Matlab's window-based FIR filter with 625 points (equivalent to 3 times sampling rate/ the lower edge of the bandpass range). The phase angles for both time series were obtained from the result of the Hilbert transform. In the accompanying code, we provide routines on calculating inter-trial coherence (a measure of potential phase resetting) and phase locking between signals that can be applied to EEG hyperscanning data. In this example, PLV estimates were computed at each time point over trials (which gives the highest degree of temporal precision) though we also provide routines for calculating these same measures in a sliding window over time within trials.

2.6.5. Simulation for event locked changes in sequential phase entrainment

To illustrate how event-locked neural responses might give rise to changes in sequential phase-based IBE (see section 2.3.1.2 of main text) two-time series were generated as sine waves at 7Hz plus white noise (sd=0.5). For time series x, we simulated 100 trials of non-phase locked activity (e.g., by adding a random phase offset, sampled from the entire 0-2pi distribution). At time 0 we simulated a phase reset, whereby the phase was abruptly shifted. Following the phase reset (e.g., t=0 for signal x) the signals became phase-locked (e.g. with an added phase offset sampled from only half of the full 0-2pi distribution). Signal y was generated in the same way as signal x, but the 'event'/ phase reset was simulated 200ms after the phase reset in signal x. Phase angles again were obtained from the result of the Hilbert transform. Phase transfer entropy was computed in a 200ms sliding window.

2.6.6. Simulation for non-event locked changes in concurrent phase entrainment

To illustrate how non-event-locked neural responses might give rise to changes in concurrent phase-based IBE (see section 2.3.2.2 of main text) we simulated two times series (640 samples at 256 Hz) x and y as basic sine waves which varied in peak frequency over time. Time series x was designed to simulate an infant alpha generator with a peak frequency range of 6-9Hz and y an adult alpha generator with a peak frequency range of 9-12Hz. For the simulated infant data, the phase offset added at each time point to a 6Hz sine wave was designed as the cumulative sum of a series of spline interpolated, linearly spaced numbers, *increasing* from 6 to 9 over the length of the time segment, minus the mean of the

upper and lower bounds of the frequency range. For the simulated adult data, the phase offset of the sine wave was set similarly, but a series of spline interpolated linearly spaced numbers *decreasing* from 12 to 9. We simulated 100 trials of data in this way. Phase angles again were obtained from the result of the Hilbert transform and PLV was calculated at each time point over trials as in the previous example.

2.6.7. Time-frequency decomposition

Single trail time-frequency power from the signals was derived by narrowband filtering the data between 2 and 25Hz in section 2.2.3 of main text and 2 and 40Hz in section 2.3.1.1 in main text. 23- 38 (respectively) filters were constructed, using MATLAB's FIR filter function, linearly spanning the frequency range. Each filter was constructed with an order of 3 times the sampling rate/ the lower limit of the frequency range (the same setting as used by EEGLab), and with a frequency spread of +/- 1.5 Hz, giving a 2/3 overlap in frequency between successive filters. Time-frequency power and phase was then derived from the resulting Hilbert transformed data. Power was obtained as the square of the absolute values derived from the Hilbert transform of the filtered data. For analysis involving time-frequency power, this was baseline normalised (decibel normalised) using activity in the -700 to -500ms time window and averaged over trials.

2.6.8. Side note on inter trial phase coherence

ITC measures the consistency of frequency band-specific phase angles over trials (time-locked to the response). The phase coherence value is computed according to:

$$ITC = \frac{1}{N}\left| \sum_{k=1}^{N} e^{i\phi(t,k)} \right| \quad (2),$$

where *N* is the number of trials and is the phase angles of a signal on trial *n*, at time *t*, Phase coherence values vary from 0 to 1, where 0 indicates no phase consistency across trials to 1 indicates oscillations take on identical phase values across trials (Lachaux et al., 1999; Delorme and Makeig, 2004).

Chapter 3 - Automatic classification of ICA components from infant EEG using MARA

The following chapter is a publication of an original article investigating methods for automated classification (keep or reject) of ICA components from infant EEG data (Marriott Haresign et al., 2021). Subheadings, figure placement, figure and table numbers, and citation style have been adapted to conform to the general thesis format. The supplementary materials (SM) for this publication are also presented within this chapter.

Abstract

Automated systems for identifying and removing non-neural ICA components are growing in popularity among EEG researchers of adult populations. Infant EEG data differs in many ways from adult EEG data, but there exists almost no specific system for automated classification of source components from paediatric populations. Here, we adapt one of the most popular systems for adult ICA component classification for use with infant EEG data. Our adapted classifier significantly outperformed the original adult classifier on samples of naturalistic free play EEG data recorded from 10 to 12-month-old infants, achieving agreement rates with the manual classification of over 75% across two validation studies (n=44, n=25). Additionally, we examined both classifiers' ability to remove stereotyped ocular artifact from a basic visual processing ERP dataset compared to manual ICA data cleaning. Here, the new classifier performed on level with expert manual cleaning and was again significantly better than the adult classifier at removing artifact whilst retaining a greater amount of genuine neural signal operationalised through comparing ERP activations

in time and space. Our new system (iMARA) offers developmental EEG researchers a flexible tool for automatic identification and removal of artifactual ICA components.

3.1. Introduction

The use of EEG in developmental cognitive neuroscience has led to a rich understanding of how the brain develops throughout early life. EEG has provided insights from birth into the development of skills such as face processing (e.g., Farroni, Csibra, Simion, and Johnson 2002), attention (e.g., Xie, Mallin and Richards, 2018), memory (e.g., Jones et al., 2020) and social interaction (e.g., Wass et al., 2018). It has also been pivotal in identifying risk factors associated with developmental disorders (e.g., Orekhova et al., 2014) and later emerging psychopathology (e.g., Jones and Johnson, 2017). However, the field is challenged by a lack of scalable, standardised tools for artifact correction. In this paper, we present one 'lossless' approach for artifact correction tuned for infant EEG data.

3.1.1. Traditional approaches to artifact removal

Despite its value, EEG recorded from paediatric populations is particularly susceptible to artifact contamination. Furthermore, it typically contains fewer sections of clean uninterrupted data due to lower recording tolerances (Gabard-Durham et al., 2018; Debnath et al., 2020). One common approach to combat this is to manually remove sections of the continuous data contaminated with artifact. However, this method of data cleaning can be problematic. For example, artifact correction in large EEG datasets can be very time consuming, and as developmental neuroscience is growing and EEG datasets are becoming larger, automated pre-processing tools are needed to efficiently process large-scale data,

taking less time than manual cleaning (Webb et al., 2015). Further manual cleaning is inherently subjective and there exist few comprehensive reviews to guide researchers (e.g., Chaumon et al., 2015). Recent studies have introduced methods for automatically identifying and removing segments of data contaminated by artifact in paediatric populations (e.g., Gabard-Durnham et al., 2018). These types of studies address the need for standardisation and speed but often rely on complete removal of artifact-affected segments. Further, many of the currently available methods for paediatric EEG have procedures designed specifically for higher electrode density recordings, therefore it is also necessary to develop artifact correction approaches that are also flexible to low-density recordings, which are often used in infant EEG studies.

Recently, there has been a drive towards the use of more naturalistic paradigms in EEG research (Risko et al., 2016; Wass et al., 2020; Holleman et al., 2020). However, naturalistic EEG recordings provide additional analytical challenges over traditional screen-based tasks. For example, in traditional screen-based/ event-related tasks in which the child is passively exposed to a set of stimuli, artifacts are more randomly distributed with respect to simulation. Removal of sections containing significant artifact can in this context be potentially beneficial, as visual experience during these sections might also be different (e.g., at its simplest the child might be fussing and not be attending to the image on the screen). However, in naturalistic paradigms, removal of whole sections of data is particularly problematic because data segments contaminated by artifact often covary with cognitive/ attentional processes of interest. Specifically, in naturalistic paradigms, the 'simulation' is often child-controlled (e.g., the child turning to the parent in a naturalistic interaction), and so artifacts are more likely to be time-locked to neural signals of interest; the removal of artifact

is thus likely to also affect the analysis of neural signals. Thus, we need approaches to the correction of artifact that remove artifactual signals from the EEG recording throughout the session, rather than removing whole segments of both signal and noise – so-called lossless pipelines.

3.1.2. Lossless approaches

Independent components analysis (ICA) is an alternative method that can be used to remove artifact from EEG data. When applied to EEG data, ICA separates the contributing sources to the scalp EEG into additive subcomponents, with varying contributions to the overall signal (Rutledge and Bouveresse, 2013; Makeig et al., 1996). Each ICA component typically contains a varying mix of neural and artifactual signals. Consideration of each component's time-frequency and topographical properties forms the basis of manual ICA classification (e.g., Chaumon et al., 2015; see also appendix B and SM Fig.1), which is typically used to separate the ICA components into two groups; components containing mostly artifactual signals and components containing mostly neural signals. As ICA itself is not a perfect method, in practice each component typically contains a varying amount of neural and artifactual signals. Some components can be clearly and easily identified manually as containing predominantly artifactual signals, whereas in other cases the mix of neural and artifactual signals is less clear and manual classification of these components becomes more subjective and based on the user's experience level.

ICA used in this way as a data preparation tool is often favoured by researchers because it allows them to subtract/ remove unwanted components (e.g., those associated with artifact)

from the EEG data without reducing the overall amount of data (hence is lossless). This is a major advantage when compared to, for example, using amplitude thresholds to remove entire sections/trials of data that are contaminated by artifact. This is particularly true for naturalistic paradigms, for the reasons given above.

We note only one other attempt to provide a system for automatic ICA classification appropriate for paediatric EEG data. The adjusted-ADJUST system (Leach et al., 2020) provides developmental researchers with an excellent framework for automatic ICA classification from typical repeated stimulus EEG data. Leach and colleagues' system achieved classification agreement with human coders of >85% with EEG recorded from 6-month-old infants. Whilst this is an impressive system, it is limited in some ways in which iMARA is not. Firstly, the adjusted-ADJUST program is set up to primarily deal with stereotypical eye movement artifact. Three of the five categories it sorts ICA components into are related to ocular motor activity. iMARA was trained on over 600 ICA components, including a wide variety of stereotyped and non-stereotyped artifacts, and so is potentially more generalisable to a wider range of artifacts. Second, adjusted-ADJUST is designed for event-locked paradigms with a repeated stimulus and is not able to incorporate EEG data from continuous/ non-event locked paradigms, which are frequently used within developmental research (e.g., to study neural entrainment in parent-infant interactions (Wass et al., 2020),), whereas iMARA is flexible to data in either format. Overall, both systems perform well and depending on the data/situation one might be more optimal than the other.

3.1.3. The MARA classification system

Many researchers manually identify which ICA-components are associated with genuine neural activity, and which are artifact. Recently, however, there have been attempts to automate this process. In this paper, we focus on one **automated** method, the Multiple Artifact Rejection Algorithm (MARA) (Winkler et al., 2011). The MARA classification system is grounded in the use of a binary linear classifier, following:

$$H = \text{sign}\,(w \cdot x + b)\ \{-1,1\} \qquad (1),$$

Where w is a weight vector obtained from samples of labelled training data, x is a feature vector containing the values of all the different component features (as illustrated in Figure 1) and b is a bias term. In short, this identifies which group the input data belong to. In the context of the MARA system, it classifies ICA components as either belonging to the 'neural' or 'artifact' group.

Classification also depends on the training data that is used. The MARA classifier was originally trained using 690 ICA-components (from an adult EEG reaction time study (n = 23 datasets)), which were manually classified as either 'neural' or 'artifact. The accuracy of the classifier was then tested on 1080 additional components from the same study. Accuracy was tested by comparing the results of the automated ICA classification to manual ICA classification. The system achieved agreement rates of approximately 91%, (i.e., 9% of components were classified differently when comparing the automated and manual classification). Accuracy was then further tested on new data from two other studies; an auditory event-related potential (ERP) paradigm (n=18 datasets); and a motor imagery BCI

paradigm (n = 80 datasets), both with different channel setups and participants. Testing the performance of the classifier on the additional data revealed agreement/error rates between the automatic and manual classification of 85/15% (Winkler et al., 2011).

Despite its popularity within adult EEG research, MARA has not received much attention within paediatric EEG research. This is perhaps because ICA itself is not widely used within traditional paediatric ERP research as a pre-processing tool. One previous study quantified the performance of MARA with paediatric EEG data. Gabard-Durnham and colleagues incorporated the classifier as part of their pre-processing tool kit (HAPPE) (Gabard-Durnham et al., 2018), applying it to samples of high density (128 channels) resting-state EEG from infants and children aged 3-36 months. The authors found that when MARA was used in conjunction with 'non-standard' approaches (e.g., wavelet thresholding of the ICA), it rejected 42% of components, but when used as part of a 'standard' pre-processing pipeline e.g., including referencing, filtering, channel rejection/interpolation, trial/ continuous data rejection and omitting the wavelet thresholding step, MARA rejected over 85% of the components. These high rejection rates highlight the importance of retraining MARA with infant data – as, typically, researchers minimise the number of components rejected to preserve as much of the original data as possible. In the present study, we aim to address the need for systems for automatic ICA cleaning of infant EEG data that can be incorporated among other standard pre-processing procedures.

3.1.4. The need to tune artifact-removal approaches to infant EEG data

Infant EEG has unique properties, requiring the design of specific tools for processing. EEG recorded from infants differs from that of children (Lepage et al., 2006) and adults

(Strogenova et al., 1999). For example, the canonical frequency bands e.g., delta (1-4Hz), theta (4-8Hz), alpha (9-13Hz), etc observed in adult EEG are observed at lower frequencies in infant EEG (Orekhova et al., 2006). Peaks in the power density spectrum that are associated with alpha activity typically observed in the 9-13Hz range in adults can be seen clearly between 6 and 9Hz in one-year-old infants (Strogenova et al., 1999) and are lower still in younger infants (Marshall et al., 2002). We also know that infant EEG tends to show greater power at lower (<6Hz) frequencies and that during development there is an observable increase in power at higher frequencies (Marshall et al., 2002). Whilst these differences have been observed in scalp level EEG data and not at a source level, this evidence highlights differences in the distribution of power at lower frequencies and the overall composition of the 1/f power density curve for infant vs adult EEG.

There is also evidence to suggest that the topographical properties of infant EEG differ from those typical of adult EEG. For example, we know that infant alpha activity projected onto central scalp electrodes is present only in later stages of infant development, presumably accompanying advances in motor skills (Cuevas et al., 2014), although the sources of these scalp activations are yet to be identified. Further, at the source level, infant EEG is often more bilaterally symmetrical than adults (Piazza et al., 2020), although strong topographical asymmetry or localisation to a specific topographical point can be a good indication of artifactual source components (Chaumon et al., 2015). This evidence highlights that infant EEG source components do contain topographically distinct properties to those of typical adult EEG. Overall, the evidence highlights the differences in the spectral and topographical properties between adult and infant EEG both at the scalp and source level. Given how important the spectral and topographical properties are for the classification of ICA

components (e.g., Chaumon et al., 2015) it should be clear from reviewing these studies that attempting to classify infant ICA components using training data from adult EEG would lead to sub-optimal results.

3.1.5 Current study: motivation and goals

In this study, we examine the performance of MARA when applied to samples of 32-channel infant EEG data acquired during naturalistic social interactions. We then adapt the MARA system to better fit the characteristics of infant EEG data. We do this in two ways; (1) by adapting the relevant time-frequency properties derived from the ICA used in classification; (2) by retraining the base classifier using data from infant EEG recordings. From here on we refer to the retrained classifier as iMARA.

To validate the performance of iMARA, we first looked at the inter-rater agreement of ICA components between three expert hand coders. We then compared MARA and iMARA to the validated, manually labelled infant ICA components across two validation studies (classifier validation 1and 2), both using different datasets. Finally (classifier validation 3), we looked at ERP data generated using the different methods to examine in greater detail their ability to remove specific types of artifact.

3.2. Methods

3.2.1. Ethics statement

This study was approved by the Psychology Research Ethics Committee at the University of East London. Participants were given a £50 shopping voucher for taking part in the project.

3.2.2. Participants

The same experimental paradigm was used for all validation datasets, but recordings were taken from different sessions (weekly sessions 1 and 8 as part of a broader, 8-week programme of research).

Dataset 1 (Validation 1), 44 healthy (23 F, 21 M) infants participated in the study along with their mothers. Infants were aged 10-12 months (mean 10.72 months, std=1.31). Dataset 1 was taken from the infant's visit 1 data.

Dataset 2 (Validation 2), 25 healthy (12 F, 13 M) infants contributed data. Infants were aged 10-12 months (mean 12.60 months, std=1.27). Dataset 2 included the same infants with data taken from visit 8.

Dataset 3 (Validation 3), 36 healthy (17 F, 18 M) infants contributed data. Infants were aged 10-12 months (mean 10.70 months, std = 1.08). Dataset 3 is a subset of dataset 1.

3.2.3. Experimental set-up and procedure

Infants were positioned immediately in front of a table in a highchair. Adults were positioned on the opposite side of the 65cm-wide table, facing the infant. Adults were given toys to play with across a tabletop and asked to "play with their infant as they would normally do at home". Adults were also asked to lower the volume of their vocalisations to reduce the level of speech-related contamination in the EEG. Dual EEG was continuously acquired from the parents and

infants for the approx. 25 min duration of the play session. For this study, we used only the infant's EEG.

3.2.4. EEG data acquisition

EEG signals were obtained using a dual 32-channel Biosemi system (10-20 standard layout). EEG was recorded at 512 Hz with no online filtering using the Actiview software.

3.2.5. EEG artifact rejection and pre-processing

A fully automatic artifact rejection procedure was adopted, following procedures from commonly used toolboxes for EEG pre-processing in adults (Mullen, 2012; Bigdely-Shamlo, et al., 2015) and infants (Gabard-Durham et al., 2018; Debnath et al., 2020). This was composed of the following steps: first, EEG data were high-pass filtered at 1Hz (FIR filter with a Hamming window applied: order 3381 and 0.25/ 25% transition slope, passband edge of 1Hz and a cut-off frequency at -6db of 0.75Hz). Although there is debate over the appropriateness of high pass filters when measuring ERP's (see Widmann and Schröger, 2012), we aimed to obtain the best possible ICA decomposition. The parameters we used were set up following recent work (e.g., Dimigen, 2020) that examined the removal of eye movement artifacts from EEG data (from a free viewing paradigm) using ICA. Second, line noise was eliminated using the EEGLAB (Delorme and Makeig, 2004) function *clean_line.m* (Mullen, 2012). Third, the data were referenced to a robust average reference (as described in Bigdely-Shamlo et al., 2015). The robust reference was obtained by rejecting channels using the EEGLAB *clean_channels.m* function (using the default settings) and averaging the remaining channels. Fourth, noisy channels were rejected, using the EEGLAB function

clean_channels.m. The function input parameters 'correlation threshold' and 'noise threshold' (inputs one and two) were set at 0.7 and 3 respectively, all other input parameters were set at their default values. Fifth, the channels identified in the previous stage were then interpolated back, using the EEGLAB function *eeg_interp.m* (mean 3.3, std, 2.1, min 0, max 9, for dataset 1. Mean 2.2, std, 1.7, min 0, max 6 for dataset 2). In some datasets, channel interpolation reduced the overall rank of the data leading to a fewer number of components than channels as is the norm with ICA. Interpolation is commonly carried out either before or after ICA cleaning, but in general, has been shown to make little difference to the overall decomposition (Delorme and Makeig 2004). Sixth, the data were low-pass filtered at 20Hz, again using an FIR filter with a Hamming window applied identically to the high-pass filter. (In the SM we also report a comparative analysis in which data were low pass filtered at 40Hz instead of 20Hz (see SM section 3.6.1.5)). Seventh, continuous data were automatically rejected in a sliding 1 second epoch based on the percentage of channels (set here at 70% of channels) that exceed 5 standard deviations of the mean channel EEG power. For example, if more than 70% of channels in each 1-sec epoch exceed 5 times the standard deviation of the mean power for all channels then this epoch is marked for rejection. This step was applied very coarsely to remove only the very worst sections of data (where almost all channels were affected), which can arise during times when infants fuss or pull the caps. This step was applied at this point in the pipeline so that these sections of data were not inputted into the ICA. The average amount of data rejected in this way was 10% (std, 8.7%, min 0%, max 35.6%) for dataset 1 and 6% (std, 5.4%, min 0%, max 20.2%) for dataset 2. Data were then concatenated and ICAs were computed on the continuous data using the EEGLAB function *runica.m.* The mean amount of data entered the ICA was 20.5 minutes (std 4.7, min 12.9, max 29.7 (mins)) for dataset 1 and 22.3 minutes (std 4.8, min 13.5, max 32.4 (mins)) for

dataset 2. In the raw data condition, we followed the same procedure but without any ICA correction.

3.2.6. Video coding

Video recordings were made using Canon LEGRIA HF R806 camcorders recording at 50fps positioned next to the child and parent respectively. Video recordings of the play sessions were coded offline, frame by frame, at 50 fps. This equates one frame to a maximum temporal accuracy of ~20ms. Coding of the infant's gaze was performed by two independent coders. Cohen's kappa between coders was >85%, which is high (McHugh, 2012). For our ERP analysis, EEG was time-locked to the onset of gaze/ saccade offline based on the video coding using synchronized LED and TTL pulses.

3.2.7. Hand identification of components for the training set

A full description of how components were identified as containing predominantly neural or artifactual signals by human coders is given in appendix B. Briefly, components were judged first on their topography, second on their power spectrum, and third on their time course, using similar principles to those suggested for adult EEG data (e.g., Chaumon et al, 2015). Components were marked as artifact/ rejected only under the null hypothesis – which in this case is that the component is not considered to contain notable amounts of neural signal. Where a researcher was in doubt over whether a component contained predominantly neural signal we opted to retain that component.

3.2.8. Inter expert reliability

As within any classification system, performance is measured concerning a criterion representing the 'true value' or 'perfect classification'. There exists no gold standard upon which to test any classifier's performance. As manual classification is the typical approach for ICA data correction (Chaumon et al., 2015) and has been used as a platform to test automatic classification in previous studies (Winkler et al., 2011), we tested the MARA and iMARA systems performance against manual ICA classification. To validate our manual coding, we asked 3 experts to independently rate ICA-components from infant and adult EEG data (SM 1.2, Table S2). We examined whether similar levels of agreement between coders could be achieved for infant ICA components as compared to those in adult data. Results are reported in section 3.1. Previous research using automated classification methods with adult data from screen-based tasks have reported error rates for inter expert agreement levels of ~10-13% MSE (Winkler et al., 2011).

The measure of performance we use in this study is mean square error (MSE), as has been used in previous automatic classification studies (Halder et al., 2007; Winkler et al., 2011). In its simplest interpretation, MSE is a measure of the error in agreement between systems. For example, an MSE of 0.25 would indicate that the automatic and manual classifiers differed on 25% of the components examined.

3.2.9. Set-up and paradigm for validation dataset 3

In validation 3 we contrasted the different classifiers' ability to remove stereotypical artifacts from an ERP analysis. This analysis examines event-locked changes relative to infants' spontaneous gaze shifts during a free-flowing naturalistic interaction. Specifically, we

examined moments where infants shifted from looking at a puppet, held at the same height as their mother's face, c.10° from the midline (counterbalanced between left and right) to looking at their mothers face, who was always positioned directly in front of the infant. For this analysis, we extracted epochs (mean 39.4, std 12.9) from the continuous data that are time-locked (time 0) to the infants' fixation onset (saccade onset at -100ms) (mean 40.8, std 11, min 18, max 64 gaze shifts were included per participant). Evidence from co-registered EEG and eye-tracking studies using free viewing experimental paradigms has shown that when visual responses (e.g., a stimulus appearing on-screen) co-vary with eye movements (e.g., horizontal/ vertical saccades) separation of these signals is possible based on their time and topographical properties (Plöchl et al., 2012). For example, some types of eye movement artifacts e.g., vertical, and horizontal eye movement transients (i.e., only lasting ~200ms) peak at ~100ms post saccade onset and have anteriorly dominated topographies, whereas visual processing components tend to peak 100-200ms after the peak of the artifact and have occipitally dominated topographies (Plöchl et al., 2012). Based on these findings and inspection of our data time-locked to saccade onsets, we set up our comparison in validation 3 between the four cleaning methods described above as follows. For comparison of removal of eye movement artifact time-locked to saccade onset, we compared peak amplitudes of potentials over frontal pole electrodes in the time window -100 (saccade onset) to 100ms (see also figure 3 for visual representation). For comparison of retention of visual response (i.e., the neural signal of interest) we compared peak amplitudes of potentials in the 200-300ms time window over occipital electrodes. We also compared amplitudes in the 200-300ms time window over central electrodes to examine how these signals propagated across the scalp. Details of which electrodes were used in each cluster can be found in the supplementary materials section (SM section 3.6.1.3).

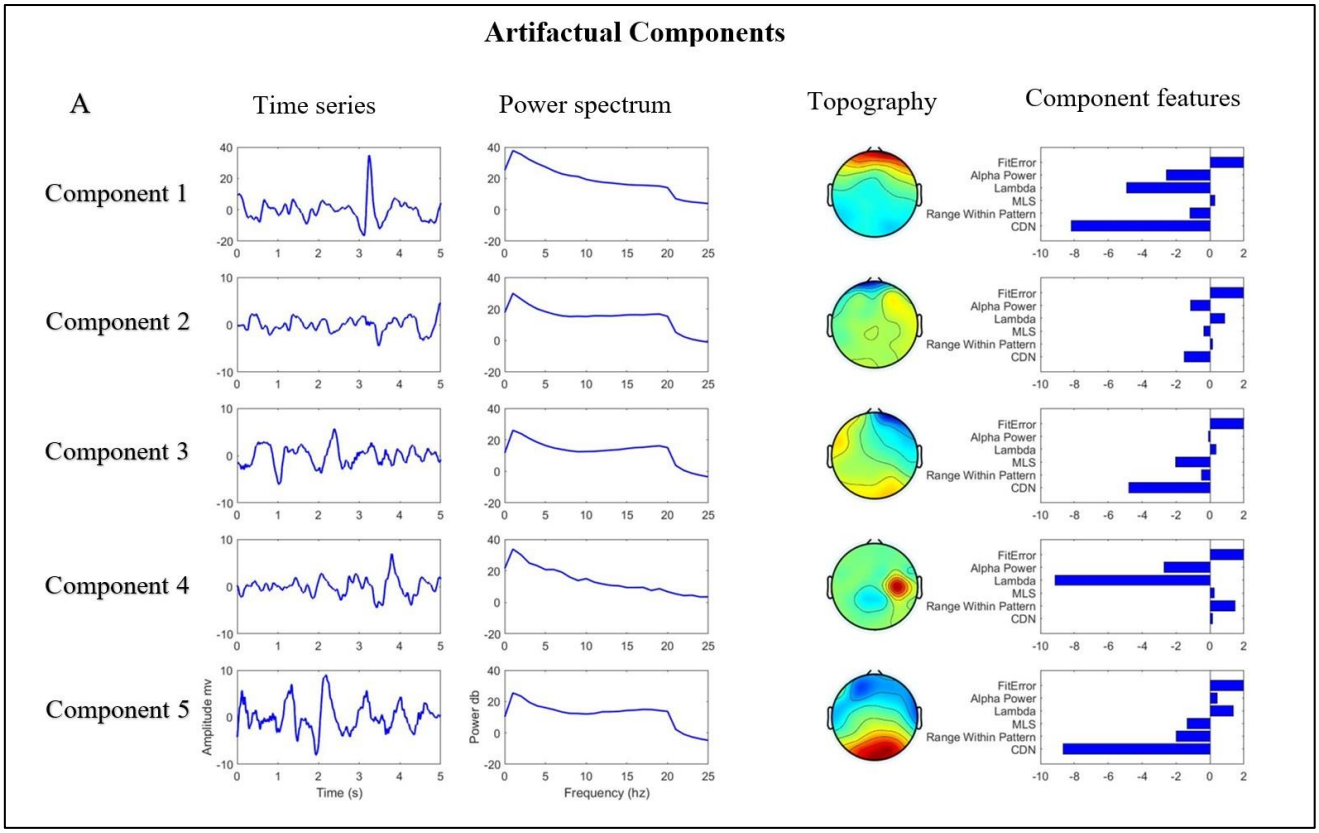
3.2.10. ERP Analysis

Differences in peak amplitude were quantified using the adaptive mean approach. This process involves identifying the peak latency of the ERP potential on a subject-by-subject basis using a broad (100ms) time window, centered around the time window of interest. For example, in our analysis, we were interested in activity in the -100 to 100ms time window. In this case, the adaptive mean approach looks for the latency of the data point with the maximum amplitude +/- 50ms around the center of the time window (0ms). Once the peak latency has been identified we took an average of the activity in a 20ms window around the peak (e.g., as described by Hoorman et al., 1998). This approach is preferred over the more basic comparison of absolute peak amplitudes which would be more susceptible to spurious noise spikes and/or unrepresentative data (Cohen, 2014). All ERP data were baseline corrected using data from the time window -1000 to -700ms pre gaze onset, based on recommendations from Cohen (2014).

3.2.11. The MARA system for automatic classification of neural/ artifactual components

The MARA classification system identifies artifactual source components from samples of EEG data. For a detailed explanation and the original source code, please refer to (https://irenne.github.io/artifacts/). In brief, Winkler, and colleagues (2011) trained a binary linear classifier to separate neural and artifactual ICA components based on a training dataset of manually labelled ICA components. The comparison between neural and artifactual components was conducted by examining six features derived from the ICA time-frequency properties (see Figure 3.1). Here we retrained the MARA system using 617 ICA components from infant EEG data taken from dataset 1 (n=25 datasets, each contributing on average 25 ICA components). We used a similar feature extraction routine as used by the original

classifier, but with a few changes to make it more specific to infant EEG data. For full details see supporting materials for chapter 3, appendix A.


Artifactual Components
A
Time series
Power spectrum
Topography
Component features
Component 1
Component 2
Component 3
Component 4
Component 5
Amplitude mv
Time (s)
Power db
Frequency (hz)
FitError
Alpha Power
Lambda
MLS
Range Within Pattern
CDN

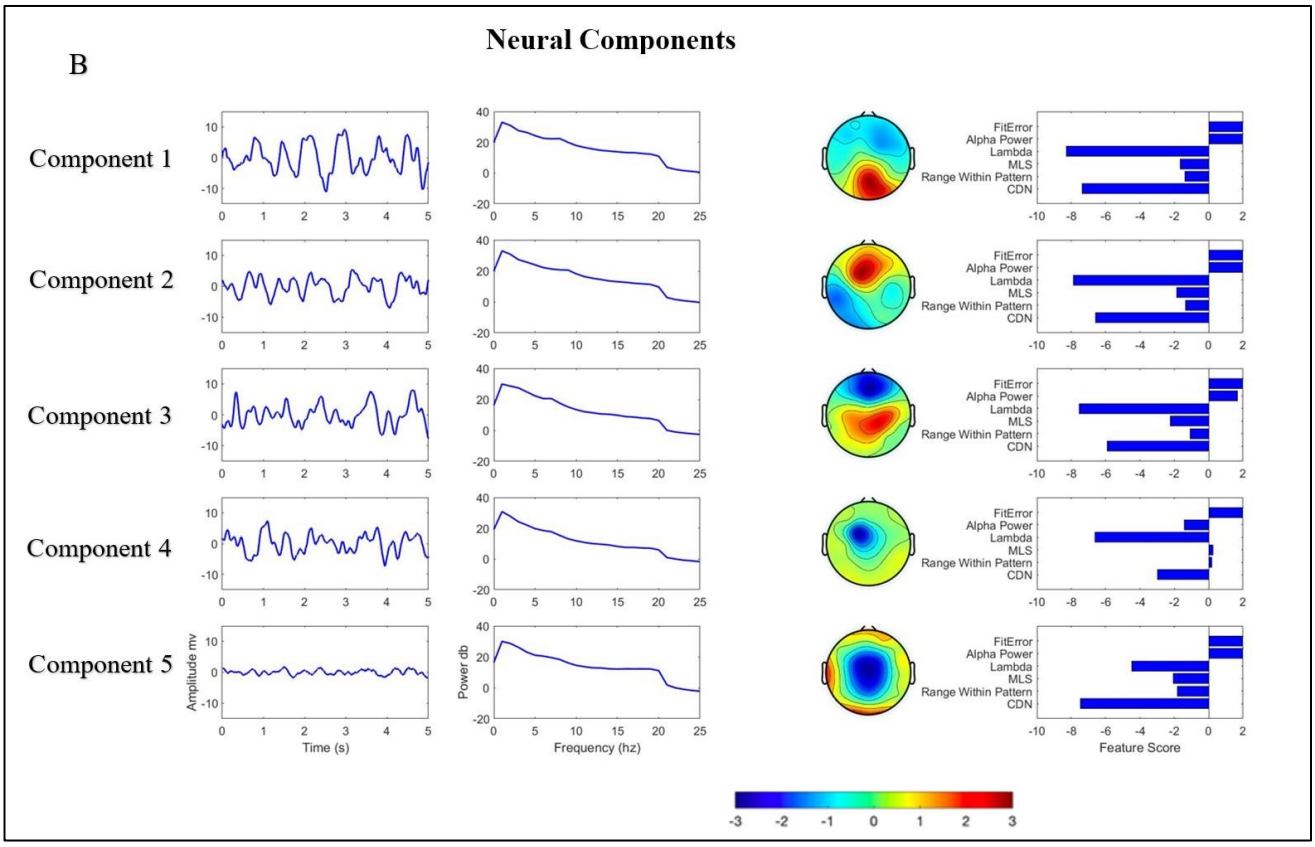

Neural Components
B
Component 1
Component 2
Component 3
Component 4
Component 5
Amplitude mv
Time (s)
Power db
Frequency (hz)
FitError
Alpha Power
Lambda
MLS
Range Within Pattern
CDN
Feature Score

Figure 3.1. Examples (taken from the present study) of neural and artifactual ICA components identified by iMARA. A) Examples of components identified as 'artifact' by iMARA. B) Examples of components identified as 'neural' by iMARA. For both, the first column shows five-second segments of the components time course; the second shows the component power spectral density; the third shows the topographical activations; and the fourth their scores for the six features used in classification. Detailed descriptions of the six features are given in Appendix A (section 3.6).

3.3. Results

First (section 3.3.1) we validated our manual classification by comparing it with manual classification from two other independent experts. Then, we perform three validation studies to test the performance of iMARA on infant data: first (classifier validation 1, section 3.3.2), we tested iMARA and MARA's agreement with manually classified ICA-components by rater 1. Second (classifier validation 2, section 3.3.3), we test iMARA and MARA's performance on ICA components from an unseen dataset. Third (classifier validation 3, section 3.3.4), we examined ERP data generated using the different systems to examine in greater detail their ability to remove specific types of artifact.

3.3.1. Inter-rater validation

To first validate our coding, we asked three experts independently to classify random subsamples of infant (n=15 datasets, average 25.6 ICs, taken from dataset 1) and adult (n=15 datasets, average 28.4 ICs, taken from dataset 1) EEG data. Full comparison details are given in SM section 1.2, Table S2. Between the 3 experts, the average disagreement rate for infant data was 18% (range across three all three experts 14-22%), whereas for adult data it was

15% (range across three experts 12-18%), which is in line with previous reports of human-human error rates for adult EEG data of 10-13% (e.g., Winkler et al., 2011). An independent sample t-test revealed no significant differences in the average agreement between adult and infant ICA-components t (14) = 0.98, p = 0.42.

3.3.2. Classifier validation 1

We tested the retrained classifier's performance against manually classified ICA components from validation dataset 1. This resulted in an averaged MSE between iMARA and the manual classification of 26.59% (sd = 9.93%, range = 54.11%). In comparison, when using the original MARA training data and the original feature extraction routine on dataset 1, the MARA classifier performed with an MSE of 38.35% (sd=15.01%, range = 60.19%). A paired samples t-test comparing the percentage of correctly identified components from validation dataset 1 for iMARA vs MARA indicated that MARA had a significantly lower level of agreement with the manual classification than iMARA t (43) = -5.94, p = <0.01. The effect size for this analysis was *d*=0.92.

3.3.3. Classifier validation 2

We then tested iMARA on an unseen dataset (dataset 2). Classification of the (645) unseen components led to an averaged MSE between iMARA and manual classification of 24.80% (std=8.22%, range=55.43%). In comparison, MARA performed with an MSE of 38.13% (std=8.12%, range=26.63%). A paired samples t-test comparing the percentage of correctly identified components from validation dataset 2 for iMARA vs MARA indicated that the original MARA had a significantly lower level of agreement with manually classified ICA components than iMARA t(24) = -4.50, p = <0.01. The effect size for this analysis was *d*=1.63.

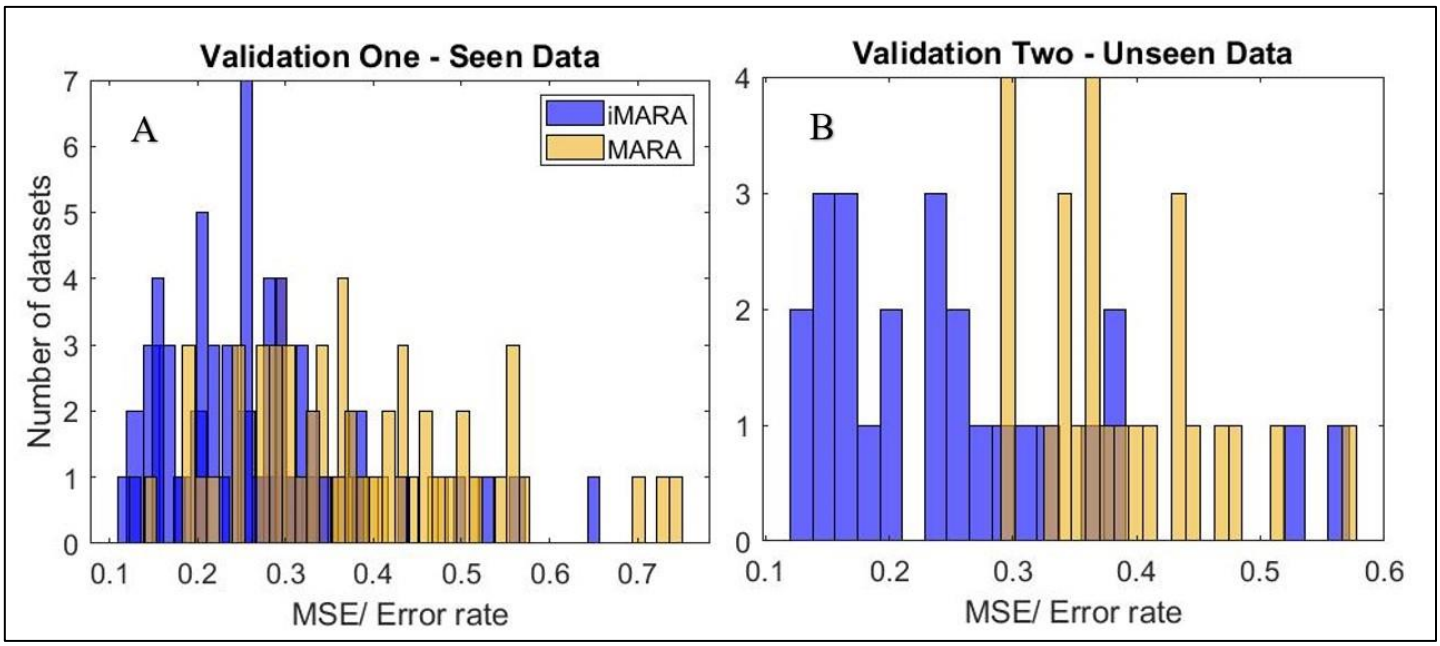


Figure 3.2. Classification performance for original (MARA) and retrained (iMARA) systems on 'seen' and 'unseen' data. A) Mean Squared Error (MSE) between original (MARA - yellow) and retrained (iMARA - blue) classifiers and manually classified ICA components for validation one (seen data) for each participant (n=44) of dataset one. B) MSE between

iMARA/MARA and the manual classification for validation two (blind data) for each participant (n=25) of dataset two.

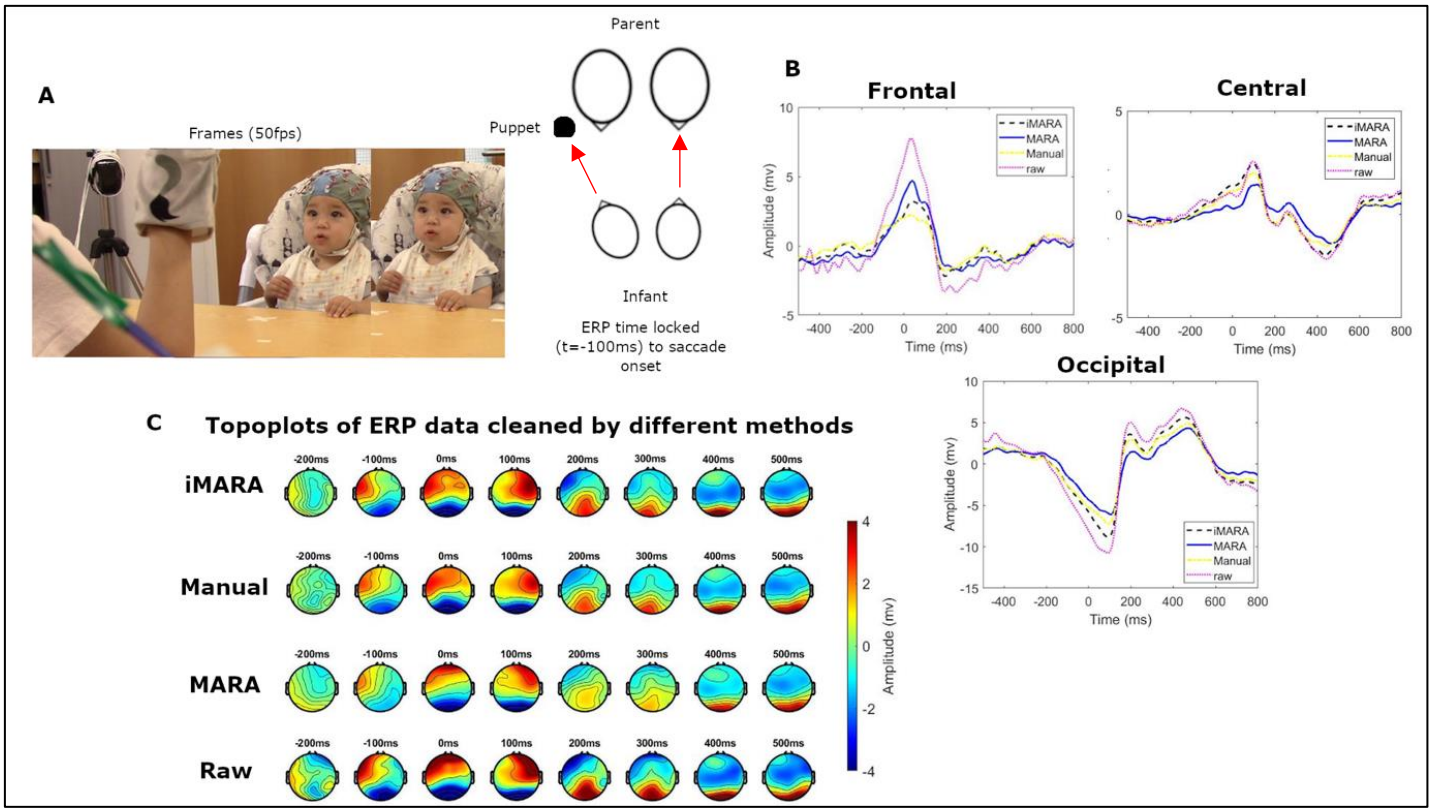


Figure 3.3. Application of different ICA classification systems to ocular artifact correction in a visual processing ERP study. A) Two-sample frames from which the time-locked gaze shift (-100ms) were identified, and a schematic showing the experimental set up in which mothers were asked to perform a puppet show with their infants. B) Grand average ERPs over frontal pole, central and occipital scalp regions. Different lines show data cleaned by the different systems, e.g., iMARA- retrained infant classifier, MARA- original classifier, Manual

classification, and uncleaned 'raw' data. C) Topoplots of ERP amplitudes, comparing the different cleaning methods to the raw data.

3.3.4. Classifier validation 3. Application to ERP study

For validation 3 (ERP analysis) we contrasted peak amplitudes (calculated on participant-level data) for each of the four methods of cleaning data (e.g., iMARA, manual cleaning, MARA and 'raw') (see Figure 3.3). In the SM section 3.6.1.7 we present a similar analysis, using time-frequency analyses rather than ERPs. We used the Tukey procedure to correct for multiple comparisons in our ERP analysis. Summary tables for all ANOVAs can be found in SM 3.6.1.1. Results from the one-way ANOVAs revealed that peak amplitudes for frontal pole ERPs in the -100 to 100ms time window were significantly lower for all ICA cleaning methods as compared to the raw un-ICA cleaned data. Peak amplitudes for iMARA were lower than for MARA, indicating that more of the ocular artifact had been removed, but this difference was not significant after correcting for multiple comparisons. For central and occipital ERPs, peak amplitudes for MARA were lower than those observed following manual cleaning and cleaning with iMARA, indicating that MARA had removed more genuine neural data. This effect was significant when examining the relationship between MARA and the raw data, but the difference between MARA and iMARA was not significant after correcting for multiple comparisons (p=.10/.11 for central/occipital).

3.4. Discussion

We retrained the popular MARA system for binary (i.e., neural or artifact) classification of ICA-components, to be more sensitive to the types of stereotypical artifacts produced during

naturalistic EEG recordings acquired from infants. Our retrained iMARA classifier classified ICA-components from samples of infant EEG with significantly greater levels of agreement with expert manual classification than the original MARA classifier. We examined how well iMARA's performance generalised to an additional blind dataset as well as its ability to remove ocular-related artifacts in a simple ERP study. Through this, we aimed to provide a tool for developmental EEG researchers wanting to implement automatic ICA cleaning.

3.4.1. Summary of retrained classifier's performance

In our first validation study, we tested MARA's and iMARA's performance against ICA-components manually classified by an expert rater on the full n=44 dataset. Here iMARA achieved a mean classification error rate of 26% (24% with outliers removed), performing significantly better than MARA (mean error rate for MARA was 38%). In the second validation, we tested iMARA on an unseen dataset, collected using the same experimental setup. In this second validation study, iMARA achieved a mean classification error rate of 25%, again significantly outperforming MARA at 38%. Overall, the differences between iMARA and MARA's agreement with the manual classification and the inter-rater agreement between humans were marginal (7-8% lower average agreement for automatic classification) relative to the overall error rates of either system (25% MSE for automatic and 18% for manual). This is consistent with the error rates between classifier-human and human-human in previous studies (e.g., 5-6% in Winkler et al., 2011). Our retrained iMARA classifier provides, therefore, a more suitable alternative for classifying paediatric ICA-components than the original MARA system. Additionally, as manual cleaning relies on a large degree of familiarity with ICA and EEG data generally, less experienced researchers using this tool can

gain insight into the types of ICA components that are commonly identified as artifacts in paediatric EEG data.

3.4.2. Application of classifiers' performance in ERP study

We also compared the performance of the iMARA and MARA to manual classification in a simple ERP study. We examined how well each classifier was able to clean the ERP data, focusing on the removal of activity over frontal pole electrodes at the onset of a saccade (gaze shift) and activity over occipital electrodes after a gaze fixation. Our analysis indicated that all methods of ICA cleaning removed statistically similar amounts of frontal pole activity from the raw (un-ICA-cleaned) data, but that neither the data cleaned manually nor iMARA removed all of the frontal pole activity associated with the eye movement artifact. This is consistent with previous research on adults, which found that standard ICA cleaning methods do not entirely remove all frontal EEG activity associated with eye movement artifacts (Plöchl et al., 2012). This is an important point which should be borne in mind in interpreting the results of EEG studies.

Results of validation 3 also show that the post-fixation (gaze onset) visual responses (ERPs) were lower in data cleaned using MARA than for the other types of cleaning, indicating that, while the original MARA classifier did successfully remove comparable amounts of the ocular artifact, it also removed significant amounts of the visually evoked potential (neural signal of interest). This is supported by further analyses (see SM section 3.6.1.4) which showed that on average MARA removed 64% of components compared to iMARA which removed 39% suggesting that MARA removed more of the total EEG variance. This effect was observed less strongly in the iMARA group, indicating that iMARA had retained more

of the original signal than MARA, but this effect was not significant after correcting for multiple comparisons.

3.4.3. Limitations of the current study

There are two explanations for the higher error rates obtained in this study, compared with the performance of the MARA classifier in the original paper, which was based on adult data (Winkler et al, 2011). First, the classification of ICA components is notably poorer when applied to lower density electrode montages. In a follow-up study, Winkler and colleagues (2014) found using the original MARA classifier that classification error rates increased from 9% to 32% when comparing 104 to 16 channel electrode setups (although for 32 channel setups it was still comparably lower ~13%) (see also SM section 3.6.1.6). This is likely due to the worsening performance of the current density norm feature (a feature estimated from the topographical properties of the data which indicates the source of the activity of the component, see appendix A for further details) with lower density setups as this feature relies on estimations of source activity and use of algorithms that are generally only recommended and applied on higher (>64) density electrode setups.

The second reason for the poorer performance compared to previous applications could be due to the increased ambiguity when classifying ICA-components from infant compared to adult EEG. This may be one of the reasons why ICA is not as widely applied within paediatric EEG research as it is within adult EEG research. In our data, we found that averaged across multiple independent coders, infant source components could only be classified with an inter-coder error rate of 18%, compared with 15% for adult data. Similar rates were also achieved when we asked the same coder (coder 1) to classify the same samples of ICA-components at a later time point. Here the agreement between coder 1 (first

and second time rating the same 384 infant ICA components) was 17%. Therefore, we suggest that ICA-components from infant EEG (particularly recorded using naturalistic paradigms) are fundamentally more ambiguous because they are more likely to contain a mixture of neural and artifactual signals, and thus are more difficult to classify binarily. Components that contain a mix of neural and artifactual signals are also likely one of the main contributing factors for why iMARA 'mislabelled' some components. We examined whether there were commonalities in the types of components mislabelled by iMARA (see SM 3.6.1.8). From visual inspection it didn't appear that iMARA was systematically mislabelling certain ICA components (i.e., with particular time-frequency and topographical characteristics) as either 'neural' or 'artifact', compared to the manual labelling. We did observe that as expected the types of components that were commonly being mislabelled (in both directions) tended to contain a mix of neural and artifactual signals and so were ambiguous even the expert human coders.

One limitation of the iMARA system is that the training data used was low pass filtered at 20Hz and so does not include examples of artfiact contaminated data beyond 20Hz. As the original MARA system was trained on adult data low pass filtered at 40Hz we performed additional analysis to examine the performance of MARA on a 40Hz filter infant EEG dataset (see SM 3.6.1.5). For this dataset, we found that both MARA and the human labelling classified over 90% of the components as artifiact, and whilst the agreement between MARA and the human labelling was fairly high (23%) it should be clear that any method that is removing over 90% of the total variance it is also removing large amounts of genuine neural activity. The high rejection rates for manual and automatic classification here are likely due to poor ICA decompositions. This is likely the result of increased muscle artifact contamination, which we know entirely overlaps with the EEG activity in the (~20-300 Hz)

spectral range (Muthukumaraswamy, 2013). Here to compensate for this and improve the ICA decomposition we have low filtered the data between 1 and 20Hz and subsequently limited iMARA's performance on EEG data beyond this frequency range. Future research may want to add additional components to the existing iMARA training dataset that include examples of components with activity at higher (20+) frequencies, assuming that these ICA decompositions can be better optimised with infant EEG.

3.4.4. Recommendations for future research

Future research might explore the iMARA's ability to separate neural and artifactual signals at different frequencies. For example, In SM 3.6.1.7 we explore the time-frequency properties of the ERP-responses shown in classifier validation 3. From these plots, it is clear that both classifiers are removing (with varying success) signal that is broadband (i.e., not frequency specific). This may be interesting for future research to explore as eye movements are commonly characterised in time or topographically but are less often characterised in time-frequency space. Having a clear picture of how ocular artifact in naturalistic data manifests in time-frequency space, as well as, having appropriate tools to identify/ remove it will be of high value to the field going forward. Additionally, it might be useful for future research to integrate iMARA as part of a fully automated EEG pre-processing pipeline either especially for paediatric EEG data or one that is flexible to adult and/or paediatric EEG data.

3.5. Conclusion

This paper presents an automatic ICA classification tool that was specifically tailored to work with infant EEG datasets and EEG data collected during naturalistic parent-infant interactions. We show that the retrained iMARA classifier achieved low classification errors and was better at cleaning stereotypical artifact from a simple visual attention ERP study than the original MARA, adult-trained classifier.

3.6. Supporting materials for chapter 3

Appendix A

Feature selection used in classification

In the original paper, the following six features were selected for use in the MARA system. These were originally chosen through an embedded feature selection process (e.g., integrated as part of the learning algorithm) whereby the authors obtained rankings of importance/ effectiveness of 38 different time/frequency/spatial features of the data (for more details see Winkler et al., 2011). This revealed that inclusion of additional features (beyond the six included) did not increase classification performance.

The following two features relate to the component spatial distribution:

Current Density Norm (CDN) - estimation of source position of a component concerning x,y,z spatial coordinates. This process involves dipole fitting the source components (using

functionality contained with EEGLAB) and applying an appropriate forward head model (we considered 2142 locations arranged in a 1 cm spaced 3D-grid) and seeking the source distribution with minimal l_2-norm (i.e., the 'simplest' solution, Winkler et al., 2014).

Components with a high CDN indicate likely artifact. For example, on Figure 1.1 a) component two, three and four all have a relatively high CDN score. These can be compared with components one, two and three of Figure 1.1b, which all have a relatively low CDN score and were classified as neural. This feature was unchanged from the original study.

Range Within Pattern - the absolute difference between the minimum and maximum of a component's pattern (spatial distribution) - i.e., how localized the activation is to one position/ electrode. Comparing components two and four in Figure 1.1a and 1.1b, we see that artifactual components have a relatively higher range within pattern indicating that these sources are more localized to a singular point, which is taken as an indication of an artifactual component to the classifier. This feature can arise, for example, from poor contact between the surface of the electrode and the scalp. This feature was unchanged from the original study.

The following two features relate to the component-time series:

Mean Local Skewness (MLS) - the mean absolute local skewness of an ICA-component time series, taken in a 1 and 15s (two separate features) sliding time window and then averaged. The idea being that blink components for example would contain epochs with very high amplitude data. This data would be more skewed than a typical alpha generator in which you would expect amplitude to be comparatively unchanged across epochs.

For example, comparing components one, two and three in Figure 1.1a and 1.1b, we see that a relatively high MLS indicates artifact, as this component's time series might contain more high amplitude noise spikes than components with a low MLS. High MLS might arise from faulty electrodes but is also an indication of an ocular motor artifact. For example, in Figure 1.1a component one, a stereotypical blink component has a relatively high MLS and contains frequent high amplitude spikes in the time series. This feature was unchanged from the original study.

The following two features relate to the component spectral distribution:

Lambda and Fit Error- the deviation of a components power spectrum from a pseudo 1/frequency curve, created by three points of the log spectrum: (1) value at 2 Hz, (2) local minimum in the band 5- 13 Hz, (3) local minimum in the band 33-39 Hz. The spectrum of muscle artifacts, characterized by unusually high values in the 20-50 Hz range, is thus approximated by a comparatively steep curve with high lambda and low fit error. Lambda and fit error are independent features; whereas lambda is a measure of the deviation from the pseudo curve just in the alpha and beta ranges (i.e., steepness of transition between the two), fit error is a measure of the deviation of the components 1/f curve from the entire pseudo 1/f curve between delta to beta.

For example from component two in Figure 1.1a and 1b, we can see that low lambda (i.e., a less steep curve between alpha and beta) indicates a neural component, whereas high lambda (i.e., a steeper upward curve between alpha and beta) indicates artifact. We can also see that fit error does not always distinguish well between neural and artifactual components in these

examples. This is because a neural component with a high alpha peak and an artifact component with a steep upward curve between alpha and beta would both give a high fit error, which can make classification using fit error alone difficult. We adjusted the frequency features to better fit the characteristics of infant EEG data. For fit error instead of taking values at 2Hz, 5-13Hz and 33-39Hz as used in MARA, we take values at 2Hz, 5-9Hz and 12-19Hz. Further for lamda instead of comparing activity in the 8-15Hz range to the pseudo 1/f curve as used in MARA, we compared activity in the 6-13Hz to the pseudo 1/f curve.

Alpha Power – The average log band power of the alpha band (8–13 Hz).

From components one and four in Figure 1.1a and 1.1b, we can see that high alpha band power indicates a neural component, whereas low alpha power indicates artifact. Instead of taking a value for alpha power in the 8-13Hz range as used in MARA, we take a value for alpha power in the 6-9Hz range.

Appendix B

ICA rejection criteria

The criteria used for manual ICA classification for infant EEG data were highly like the principles suggested for adult EEG data (e.g., Chaumon et al., 2015). Components were marked as artifact/ rejected only under the null hypothesis – when the component is not

considered to contain notable amounts of the neural signal. Where a researcher was in doubt over whether a component contains real EEG (neural) we opted to retain that component.

All selection of components was performed using the interface provided by EEGLAB's *pop_selectcomps* function. Often, EEG researchers only reviewed the first ~10 components, as later components account for little overall variance. The machine-based classifier, however, reviews all components on an individual basis, and so for appropriate comparison coders were asked to rate all components on an individual basis.

Components were judged first on their topography, second on their power spectrum, and third on their time course, according to the flow chart below:

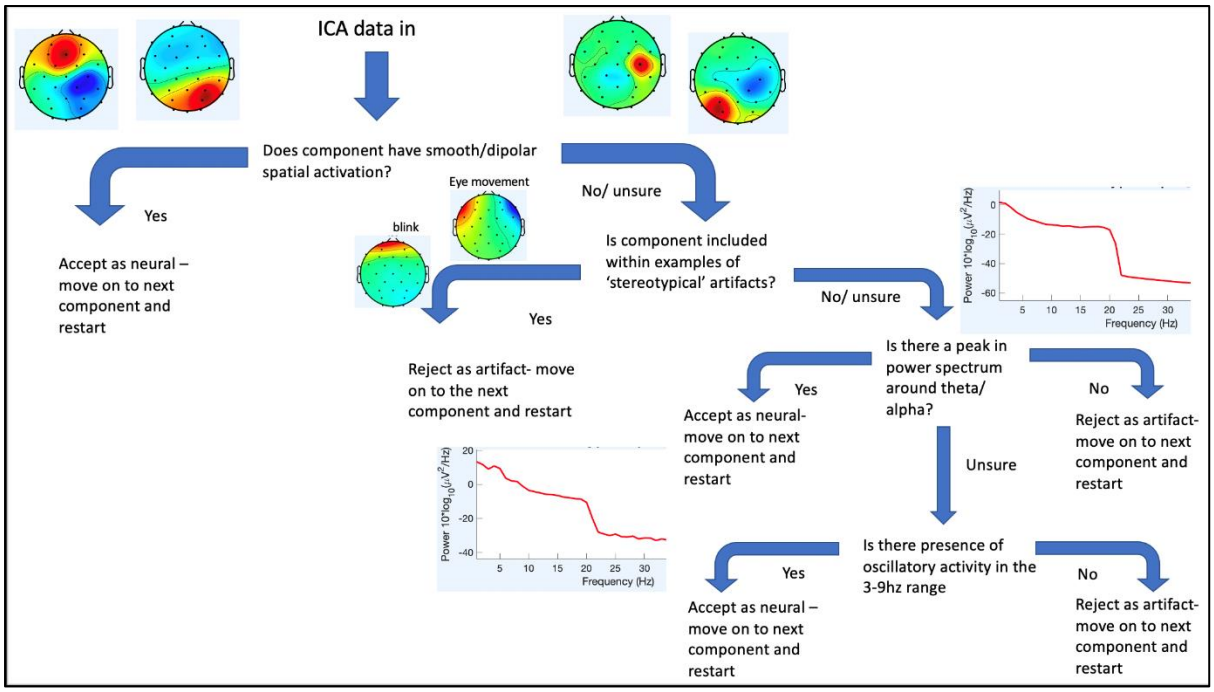
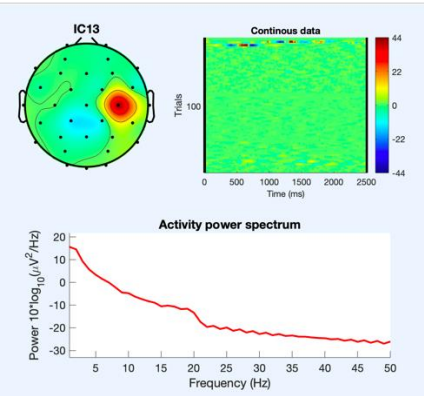


Figure 3.4. Flow diagram of human ICA classification:

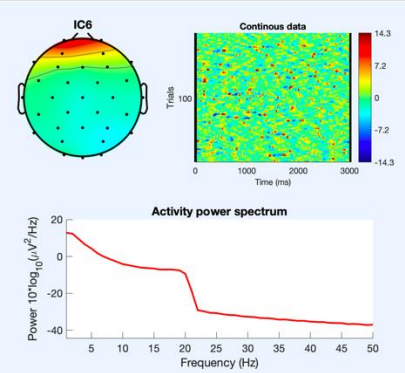
Stage 1. Evaluation of topographical activation

Neural components were largely identified by the presence of smooth, dipolar spatial activity. These components were often immediately identifiable in the *pop_selectcomps* component overview and did not need further investigation. A variety of different criteria based on topographical activation were applied:

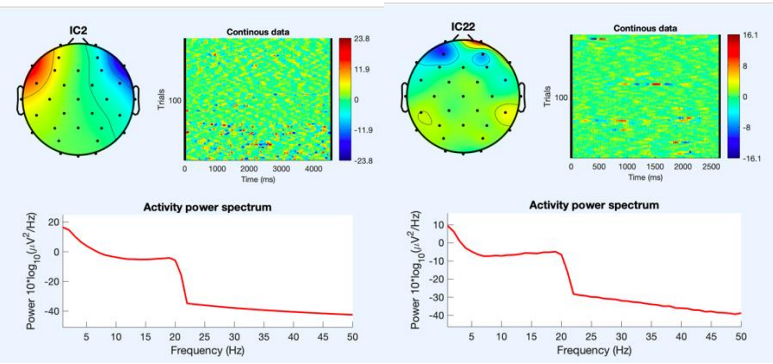
Localisation to one electrode. Components with very localised topographical (i.e., localised to one electrode) activations were always a cause for further investigation (in these cases, coders would move to stage 2). For example:



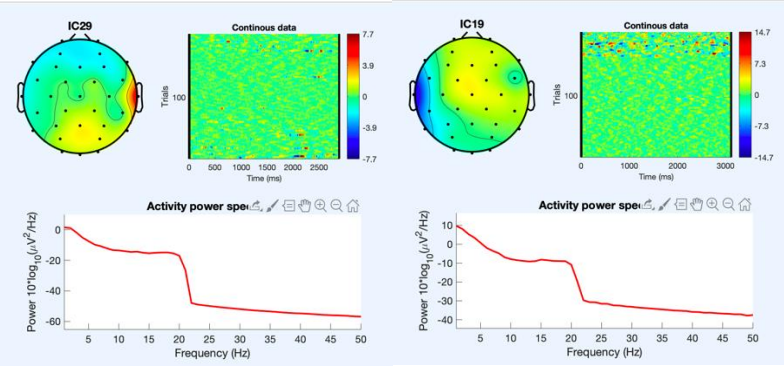
Frontal peripheral activation. Certain stereotypical artifactual components were readily identifiable from their topographical maps. For example, components with strong peripheral activations and in particular components with strong/ very localised frontal pole activation (blinks). For example:



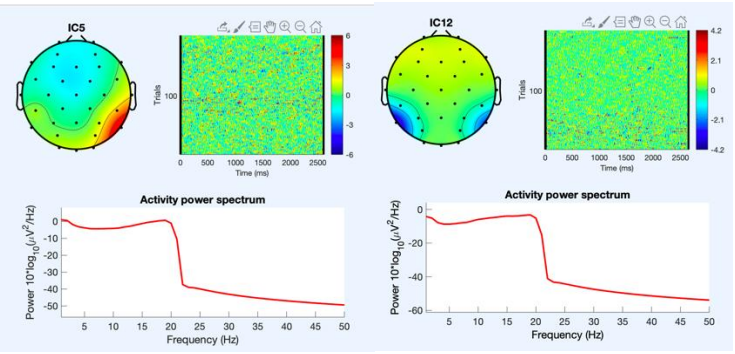
Bilateral frontal topography. Components with opposite bilateral frontal topography often indicated horizontal eye movements. For example:



Temporal peripheral activation. Components with strong/ very localised temporal activation often indicated jaw/ speech related artefact. For example:

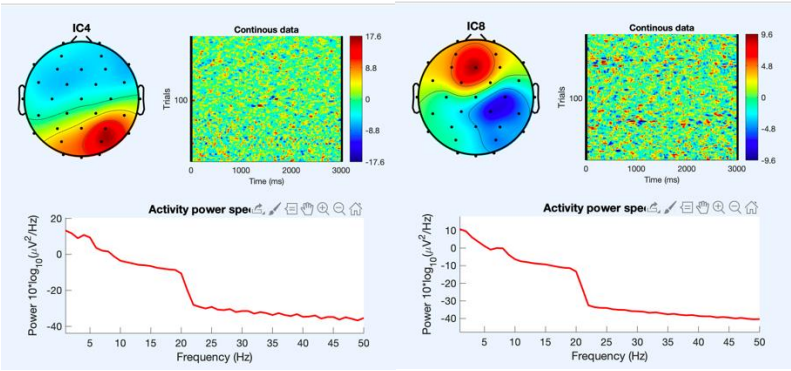


Activation around P7/P8. Components with bilateral/strong activation around P7/ P8 (on 10-20 32 channels layout) often indicated neck movement. For example:

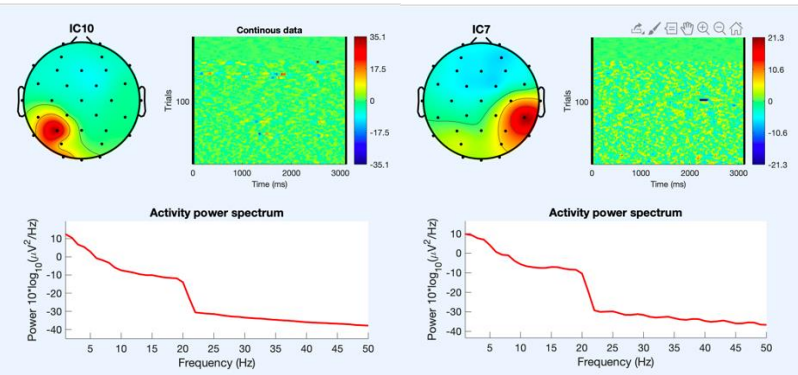


Stage 2. Evaluation of the power spectrum

Neural components were identified by 'peaks' in the power spectrum in theta (3-6Hz) and/ or alpha (6-9Hz) range. For example:



Components with more mixed neural and artefact sources with more subtle peaks in theta/alpha were also identified at this stage as neural. For example:



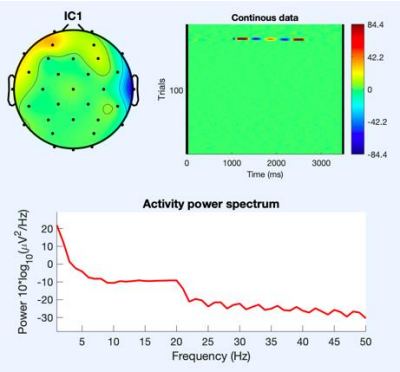
Artifactual components were predominantly identified at this stage through lack of alpha/ theta peak. Further, if a component had no clear 'peak' in alpha but accounted for a lot of total variance coders would move to stage 3 (see below).

Components with power at high frequencies (muscle/ speech artifact) were also marked as artifact at this stage (for examples see the earlier examples for speech and movement-related artifacts).

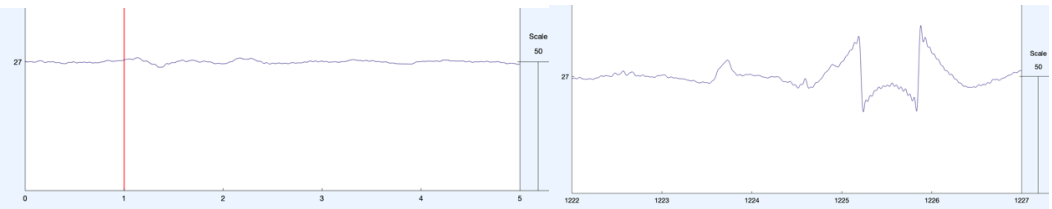
Components with activity beyond 20Hz were also excluded as the current dataset is low pass filtered at 20Hz – see example below under infrequent high amplitude noise spikes.

Stage 3. Evaluation of time course

Artifactual time courses were identified predominantly from their ERP image top right of below figure. If time course activation was mainly driven by infrequent high amplitude spikes (spots of extreme dark colour in ERP image) then it was marked as artefact. For example:

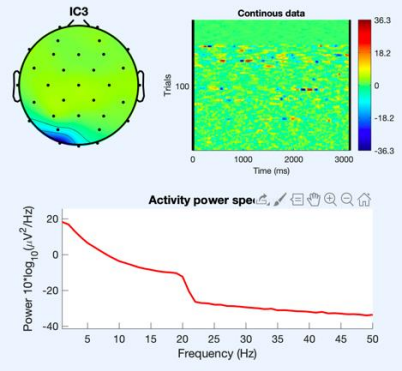


A segment of the component time course:

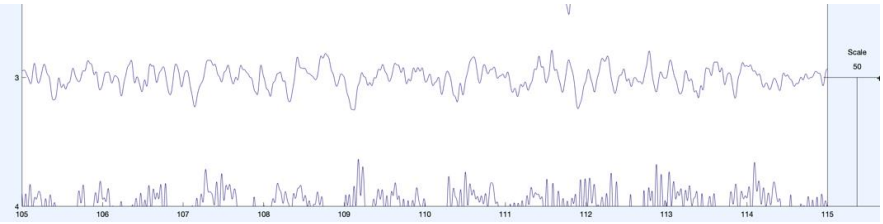


This is different from the example component below as this components time course is predominantly flat throughout.

Other components showed ‘high’ theta/alpha power but no peak. For example:



A segment of the component time course:



Supplementary Materials

3.6.1.1 Validation 3: Results of ANOVAs

Table 3.1. Summary results of One-Way ANOVA for each scalp region. For frontal pole electrodes, peak amplitudes were compared in the -100 to 100ms time window. For central and occipital components, amplitudes were compared in the 200 to 300ms time window. Electrode groupings used are shown in Table 3.1.

Group ID		'F'			'*p*'		
		Frontal	Central	Occipital	Frontal	Central	Occipital
iMARA	MARA	2.17	2.06	6.43	0.89	0.10	0.11
iMARA	Manual	3.92	1.54	4.39	0.85	0.70	0.88
iMARA	Raw	-2.39	0.70	1.88	<0.01	0.79	0.65
MARA	Manual	4.74	0.57	1.39	0.42	0.60	0.41
MARA	Raw	-1.57	-0.26	-1.12	<0.01	<0.01	<0.01
Manual	Raw	-3.32	0.26	0.91	<0.01	0.20	0.22

3.6.1.2. Inter expert reliability

Table 3.2. Error rates between expert coders on an n=15 subsample of the total data. Each cell shows agreement for infant data (left) and adult data (right)

	Rater 1	Rater 2	Rater 3
Rater 1		0.22/0.12	0.14/0.15
Rater 2	0.22/0.12		0.19/0.18

Rater 3	0.14/0.15	0.19/0.18	

3.6.1.3. Electrode positions

Table 3.3. Channel clusters and corresponding 10–20, 32-channel Biosemi positions

Clusters	10-20 Positions	Biosemi 32 channel electrodes
Frontal pole	Fp1, AF3, AF4, FP2	1, 2, 29, 30
Frontal	F7, F3, F4, F8, Fz	3, 4, 27, 28, 31
Central	C3, CP1, CP5, CP6, CP2, C4, Cz	5, 6, 8, 9, 10, 21, 22, 23, 25, 26, 32
Temporal	T7, T8	7, 24
Parietal	P7, P3, Pz, P4, P8	11, 12, 13, 19, 20
Occipital	PO3, O1, Oz, O2, PO4	14, 15, 16, 17, 18

3.6.1.4. ICA components removed by each method

Table 3.4. Average number and percentage of ICA components removed by each method

System	Mean number of components removed (% of total)
Original MARA	18 (64%)
Retrained MARA	11 (39%)

Hand cleaned	13 (46%)

3.6.1.5. 40Hz subsample replication

All the data used in this study were narrow band filtered between 1 and 20Hz. This was done before ICA correction to improve the signal to noise ratios and the quality of the ICA decomposition. However, in their tutorial (https://irenne.github.io/artifacts/) Winkler and colleagues (2011) note that application of MARA to narrow band filtered data might result in suboptimal performance, as the spectral features are calculated on the power spectrum between 2 and 39Hz. We, therefore, wanted to test the original classifier's performance on naturalistic infant data that had not been narrow band filtered. We processed a subsample of 15 datasets in the same way as described in the main text except for the data was low pass filtered at 40Hz this time instead of 20Hz. We then ran ICA, and all source components were again hand labelled by an expert. Results indicated an MSE between the classifier and human labelling of 23% indicating a good degree of similarity. However, the classifier was removing on average 97% of the approx. 32 components, i.e., retaining on average only one source component. The human labelling also removed 90% of the source components, more than double the amount as the data used in the main text. Whilst the MSE is slightly lower than the retrained classifier in the main text, it is clear that this method is far from optimal as clearly if any method is removing over 95% of the total variance it is also removing large amounts of genuine neural activity. The high rejection rates for manual and automatic classification here are likely due to poor ICA decompositions. This is likely the result of increased muscle artifact contamination, which we know entirely overlaps with the EEG activity in the (~20-300 Hz) spectral range (Muthukumaraswamy, 2013). As naturalistic infant EEG data inherently contains more muscle movement (and therefore more artifact contamination) than

screen-based paradigms, in the main text we apply a 20 Hz low pass filter to the data to attenuate some of this activity. Comparing the results of the main text with the result obtained when applying the original MARA with data filtered at 40 Hz, it is clear that the 20 Hz low pass filter is greatly improving the ICA decompositions, resulting in a much-reduced percentage of components labelled as artifact.

3.6.1.6. MARA adapted strategy

In their follow up study in which Winkler and colleagues (2014) tested the robustness of their classifier on a variety of novel experimental paradigms and electrode setups, they found that when applied to data with lower density electrode setups e.g. 16 or 32 channels, the classifier's error rate increased linearly from an MSE between automatic and manual classification from 9 to 32%. This suggests that the classifier performs significantly worse with lower density electrode setups. The reason for this decrease in performance was due to the spatial features performing notably worse (Winkler et al., 2014). The MSE of the CDN feature on 32 and 16 channel setups rose to over 50% compared to 12% with higher density setups. To improve the performance of the original MARA classifier with lower density EEG recordings, the authors proposed an adaptive strategy in which the classifier is subtly adapted to fit the study-specific electrode montage i.e., re-training the classifier on the patterns cut to the specific electrode setup. In their study application of the adaptive strategy with 16 and 32 channel setups lead to an MSE of approx. 16%. We also tested this adaptive strategy on our 32 channel infant EEG data to see whether this led to a similar error rate. We did this by adapting the electrode montage used for spatial feature identification to the Biosemi 32-channel layout but using the same training data as used by the original MARA classifier. This led to an MSE between the adapted strategy classifier and the human labelling of 43%.

Neither the original classifier nor the adapted strategy performed well when applied to our infant EEG data. On one hand, we have a system that achieved decent rates of agreement with hand labelling i.e., MSE of 23% but removes over 95% of the data (e.g., original MARA); on the other hand, the adaptive strategy removes fewer components but has a much higher error rate. Therefore, the final option is to treat the infant data as distinct, retraining the classifier using infant source components and the most salient features for infant EEG.

3.6.1.7. Time-Frequency analysis of ERP data presented in validation study 3

To further investigate the removal of ocular artifact reported in validation 3. We examined time-frequency responses for the different methods to assess the time-frequency representations of the signal and how this was affected by the ICA cleaning methods. For this analysis single-trial data were first decomposed into their time-frequency representation by multiplying the power spectrum of the EEG (obtained from the fast-Fourier-transform) by the power spectrum of complex Morlet wavelets [$e^{i2\pi tf}e^{-t^2/(2\sigma^2)}$, where t is time, f is the frequency (which increased from 2 to 16 Hz in 15 linearly spaced steps), and σ defines the width of each frequency band, set according to $\sigma/(2\pi f)$, σ was set to scale with increases in the centre frequency of the wavelet. We set this parameter to increase logarithmically from 3-10 cycles in 15 increments], and then taking the inverse fast-Fourier-transform. From the resulting complex signal, an estimate of frequency band-specific power at each time point was defined as the squared magnitude of the result of the convolution Z (real$[z(t)]^2$ + imag$[z(t)]^2$). Power was then decibel normalised to data in the -1000 to -700ms time window.

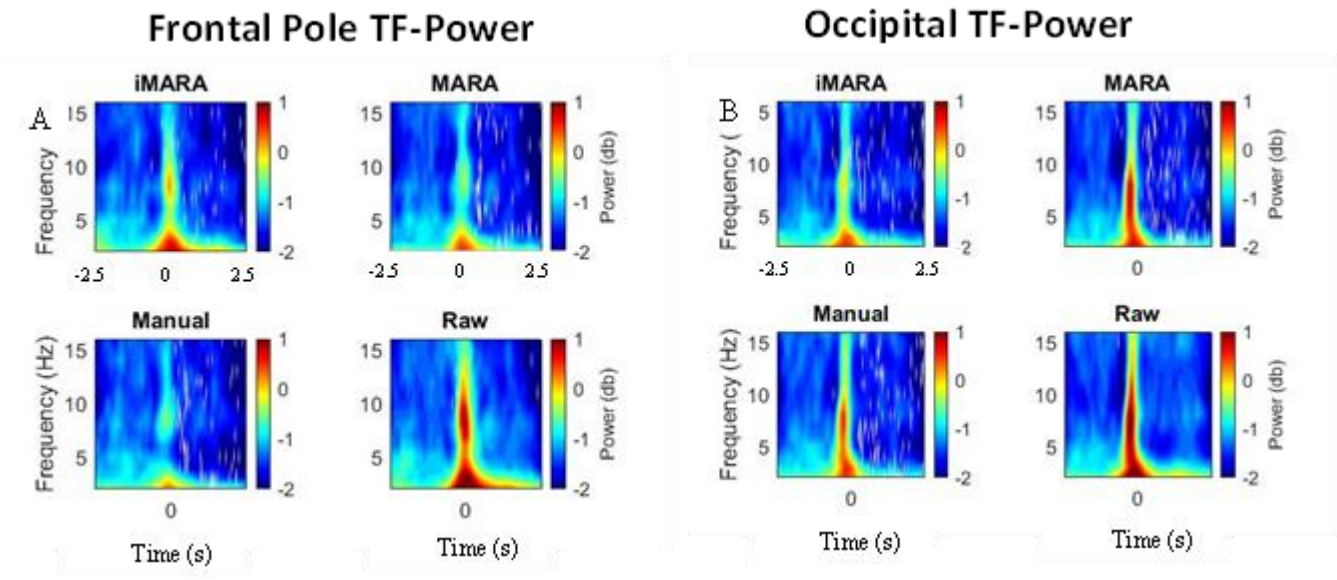


Figure 3.5. Time-frequency power relative to the onset of infant looks to partner in data cleaned using original 'MARA', retrained 'iMARA' and manual ICA classification and compared to raw data. A) shows time-frequency plots for data cleaned using 4 methods over frontal pole electrodes. B) shows time-frequency plots for data cleaned using 4 methods over occipital electrodes

From Figure 3.5 we can see that all methods of cleaning resulted in broadband removal of the signal both frontally and occipitally. We can also see that the different methods are to a varying degree removing most but not all the ocular artifact time-locked to the shift in attention. This may be an interesting possibility for future research to explore as commonly eye movements are characterised in time or space but are less often characterised in time-frequency space.

3.6.1.8. Components mislabelled by iMARA

We performed an additional visual inspection of components that were mislabelled by iMARA to see if they shared any common characteristics. We did this by first looking at components that were labelled as artifact by iMARA, but neural by manual labelling and second looking at components that were labelled as artifact by the manual labelling, but that were classified as neural by iMARA. From visual inspection, we did not observe any strong patterns that distinguished between these two subcategories, i.e., it didn't appear that iMARA was systematically failing to reject ICA components with certain time-frequency or spatial properties. One thing that we did, however, observe from components mislabelled as artifact by iMARA (as compared to the human labelling) was that they tended to have a 'high' range within pattern (RWP). As we have described in appendix A, a high RWP generally indicates artifact to iMARA. From the examples below (Fig. 3.6) of components mislabelled as artifact by iMARA we can see that all these components have a 'high' RWP, but that they also contain some evidence of a real neural signal in the EEG spectra (visual peaks in 1/f curve at theta/ alpha). As the amount of theta/ alpha power is low and the RWP is 'high', iMARA classified these components as artifact, but as there is some (albeit subtle) evidence for neural signal the human labelling classified them as neural. We also observed that in some cases in which the component's time-frequency and spatial properties were completely distorted (see Fig. 3.7 for examples) likely due to poor connection between the electrode and the scalp, the 'high' amount of relative alpha power resulted in iMARA mislabelling these components as neural. This is because 'high' alpha power (see appendix A) generally indicated a neural component to iMARA. In future work, we plan to add more components to the training set (currently using just over 600 ICA components) so that the classification system can benefit from a larger and more diverse pool of data.

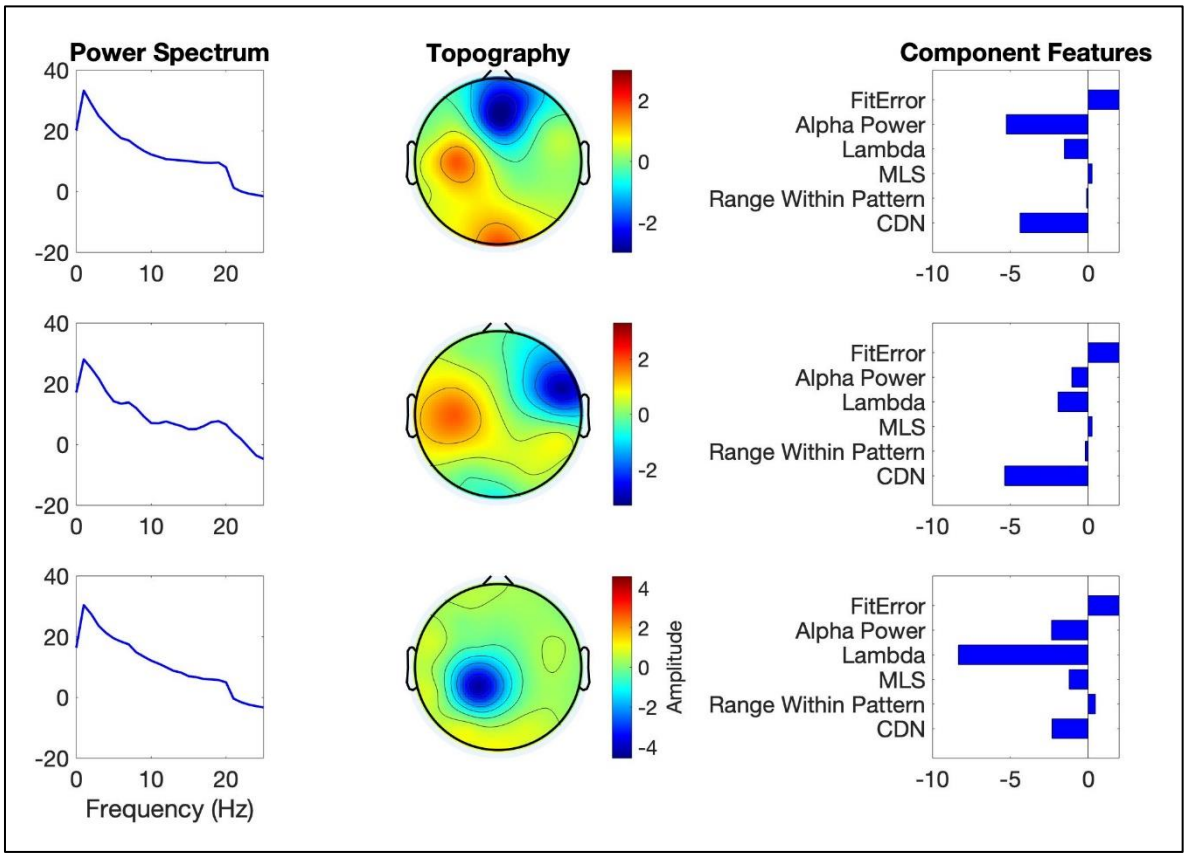


Figure 3.6. Example components 'mislabelled' as artifact by iMARA. Column one shows the component power spectral density; column two shows the topographical activations; and column three their scores for the six features used in classification. Detailed descriptions of the six features are given in Appendix A.

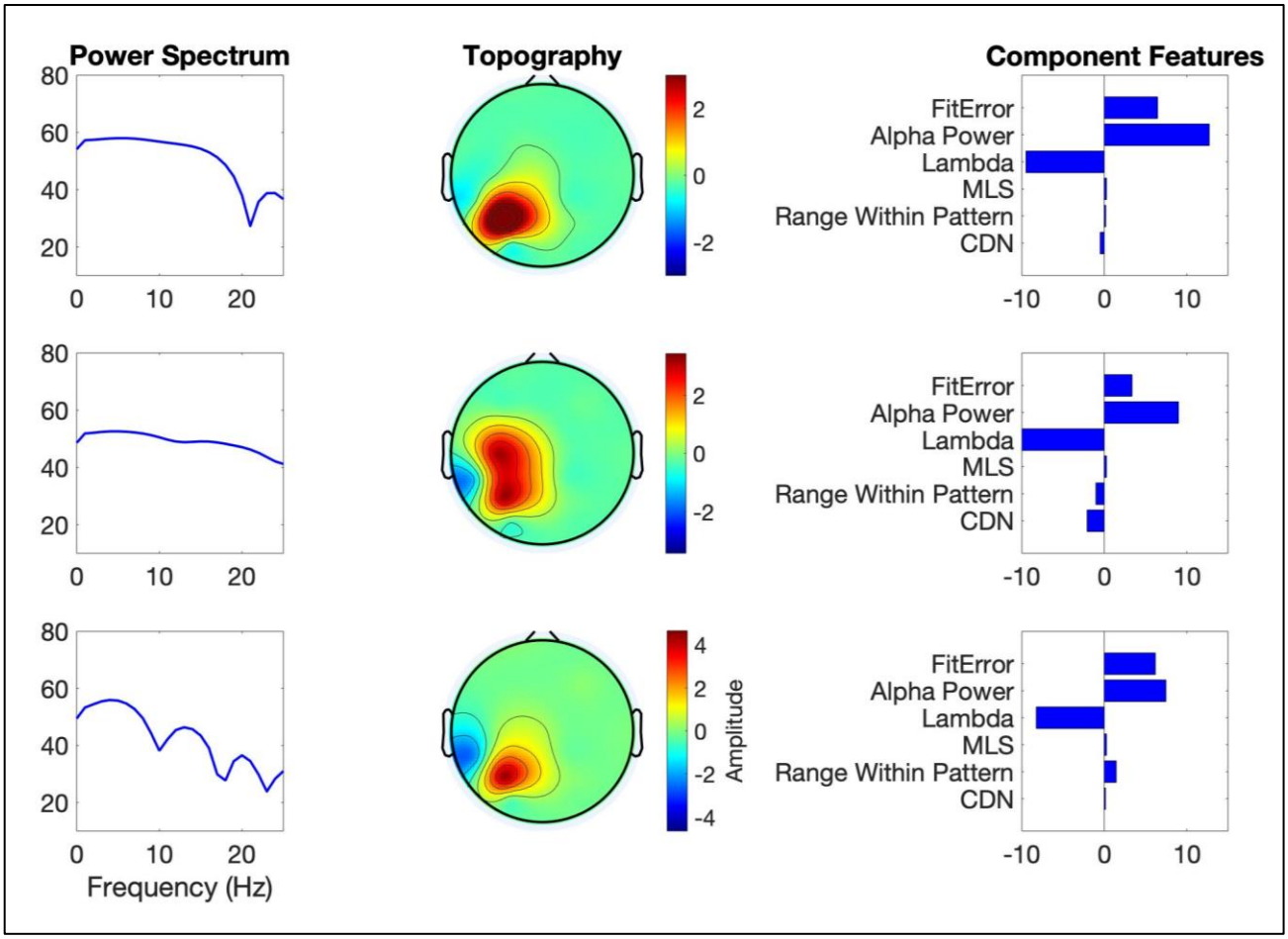


Figure 3.7. Example components 'mislabelled' as neural by iMARA. Column one shows the component power spectral density; column two shows the topographical activations; and column three their scores for the six features used in classification. Detailed descriptions of the six features are given in Appendix A.

Chapter 4 - Gaze onsets during naturalistic infant-caregiver interaction associate with 'sender' but not 'receiver' neural responses, and do not lead to changes in inter-brain synchrony.

The following chapter is a publication of an original article investigating the mechanisms that give rise to parent-infant inter-brain synchrony around moments of mutual gaze (Marriott Haresign et al., 2023). Subheadings, figure placement, figure and table numbers, and citation style have been adapted to conform to the general thesis format. The supplementary materials (SM) for this publication are also presented within this chapter.

Abstract

Temporal coordination during infant-caregiver social interaction is thought to be crucial for supporting early language acquisition and cognitive development. Despite a growing prevalence of theories suggesting that increased inter-brain synchrony associates with many key aspects of social interactions such as mutual gaze, little is known about how this arises during development. Here, we investigated the role of mutual gaze *onsets* as a potential driver of inter-brain synchrony. We extracted dual EEG activity around naturally occurring gaze onsets during infant-caregiver social interactions in N=55 dyads (mean age 12 months). We differentiated between two types of gaze onset, depending on each partners' role. 'Sender' gaze onsets were defined at a time when either the adult or the infant made a gaze

shift towards their partner at a time when their partner was either already looking at them (mutual) or not looking at them (non-mutual). 'Receiver' gaze onsets were defined at a time when their partner made a gaze shift towards them at a time when either the adult or the infant was already looking at their partner (mutual) or not (non-mutual). Contrary to our hypothesis we found that, during a naturalistic interaction, both mutual and non-mutual gaze onsets were associated with changes in the sender, but not the receiver's brain activity and were not associated with increases in inter-brain synchrony above baseline. Further, we found that mutual, compared to non-mutual gaze onsets were not associated with increased inter brain synchrony. Overall, our results suggest that the effects of mutual gaze are strongest at the intra-brain level, in the 'sender' but not the 'receiver' of the mutual gaze.

4.1. Introduction

Most of our early life is spent in the presence of an adult social partner. Most early attention – and, in particular, most early cognitive learning – takes place in social settings (Csibra and Gergely, 2009). But almost all of our understanding of how the brain subserves attention and learning has come from studies that measure individual brains in isolation.

In recent years our understanding of early social and communicative development has relied heavily on studying ostensive signals, defined as signals from a communicator to generate an interpretation of communicative intention in an addressee. It has been argued that, from shortly after birth, infants' brains are sensitive to ostensive signals (such as direct gaze, smiles and infant-directed vocalisations), and that 'sender' communicative signals play a key role in supporting early learning exchanges (Werker and Yeung, 2005; Csibra and Gergely, 2009; Csibra, 2010; Southgate et al., 2010; Begus et al., 2014; 2016; Ferguson & Lew-

Williams, 2016). In this paper we focus on mutual gaze, which is a widely studied ostensive signal.

Farroni et al., 2002 found that images of faces showing direct vs averted eye contact elicited greater amplitude event related potentials (ERPs) in infants even 2- to 5-days after birth. Grossman et al., 2007 observed greater gamma power activation in 4-month-old infants in response to facial stimuli with direct gaze vs averted gaze, and where an experimenter engages in mutual gaze before looking towards an object, infants show enhanced neural processing of those objects (Parise et al., 2008; Hoehl et al., 2014). Of note, findings of gaze orientation on ERP amplitudes have not replicated well in developmental research; for example, Elsabbagh and colleagues (2009) found no significant effect of gaze type, comparing ERP amplitudes (see control group comparison); and findings are also largely mixed in adults (e.g., Watanabe et al., 2001; Taylor et al., 2001b; Watanabe et al., 2002; Itier et al., 2007; Conty et al., 2007; Ponkanen et al., 2011). Despite the inconsistencies, these findings have contributed to the popular idea that infants are highly sensitive to their partner's social signals during early learning exchanges (Hains & Muir, 1996b; Symons et al., 1998; Farroni et al., 2004; Werker & Yueng, 2005; Werker et al., 2007; Hoehl et al., 2008; see Cetincelik et al., 2021 for a review). This raises basic questions of in what contexts, and under what circumstances are infants sensitive to their partner's social signals.

One important limitation of research measuring infants' neural responses to social stimuli presented on a screen, however, is its limited ecological validity. Historically, the majority of studies into early social development have measured infants' passive responses to viewing a series of static images that flash on and off on a screen in a predetermined sequence. The real world, in contrast, is interactive, contingent and continuous. In recent years an increasing number of researchers have begun to recognise that, to study how the infant brain subserves social interaction, it is necessary to actually study it in interactive contexts (Schilbach et al.,

2013; Redcay & Shilbach, 2019; Wass et al., 2020; Wass & Goupil, 2022). Recent behavioural findings have suggested important differences between screen-based simulacra of social interaction and actual social interaction. For example, recent studies (Franchak et al., 2011; Yu & Smith, 2013) have shown that infants rarely look to their caregiver's face and eyes during free-flowing interactions, this is in contrast to what has previously been shown using screen-based tasks (e.g., Vecera & Johnson, 1995; Farroni et al., 2004). So far, the neural processing of mutual gaze has largely been investigated in un-ecological contexts, and in the absence of real social interaction. Consequently, two important questions remain unanswered; 1) is mutual gaze really a salient communicative signal during free-flowing social interactions occurring in rich, continuous, natural scenes? 2) how does intra and inter-brain activity support the processing of sender and receiver ostensive signals such as mutual gaze when both partners are engaged in a free-flowing, bidirectional exchange of information?

Mutual gaze and inter-brain synchrony

Another topic that has shown rapidly burgeoning popularity in recent years is inter-brain synchrony. At the neural level, inter-brain synchrony can be defined as a dyadic mechanism, wherein temporally coordinated patterns of brain activity between two interacting individuals supports aspects of their ongoing social interaction (Holroyd 2022). A number of studies have observed increased inter-brain synchrony during mutual gaze. The majority of this research claims to measure inter-brain synchrony, although we recognise that not all of these studies will meet the framework of inter-brain synchrony set out in more recent theoretical accounts (Holroyd, 2022). Kinreich and colleagues (2017) observed significantly correlated gamma (30-60Hz) activity between interacting adults during social interaction. Higher

interpersonal gamma correlations were also associated more strongly with mutual vs non-mutual gaze. Similarly, Luft and colleagues (2021) found that mutual gaze was associated with higher inter-brain gamma band (30-45Hz) coherence (a spectral measure based on correlation) between interacting adults than non-mutual gaze. In the developmental literature, our group investigated inter-brain synchrony in 7.5-month infant-adult dyads during moments of mutual and non-mutual gaze (Leong et al., 2017). During a live social, but not interactional condition infants observed an adult singing nursery rhymes, who was instructed to look either directly at the infant, directly at the infant with their head turned at a slight angle, or away from the infant. Consistent with research on adults, we found greater infant-adult neural synchrony during moments of mutual vs non-mutual gaze, measured using partially directed coherence (PDC-a spectral Granger causal measure of synchrony) in Theta (3-6Hz) and Alpha (6-9Hz) band activity. This study thus suggests that the impact of mutual gaze on inter-brain synchrony found in adult-adult dyads (Kinreich et al., 2017; Luft et al., 2021) is already present early on in development, though possibly in lower frequency brain rhythms.

Sender/ Receiver mechanisms of inter-brain synchrony

As recent theoretical accounts have highlighted (Burgess 2013, Hamilton, 2021; Holroyd, 2022) inter-brain synchrony can reflect underlying mechanisms of varying complexity. To date, research investigating inter-brain synchrony during social interaction has exclusively measured this using non-event locked analyses, i.e., inter-brain synchrony values are averaged across whole conditions and/ or whole interactions and not time locked to any specific events within the interaction. Previously we have argued that in order to distinguish between different mechanisms that might give rise to inter-brain synchrony it is important to use event locked analyses (Haresign et al., 2022), i.e., analyses that focus on measuring fine-grained temporal changes in inter-brain synchrony, time-locked to specific behaviours/ events

within social interactions. Additionally, when trying to differentiate inter-brain synchrony from other forms of inter-personal neural synchrony, it is also important to measure dyadic dynamics, e.g., how both partners' behavior and neural activity contribute to establishing inter-brain synchrony.

We have suggested that one leading candidate mechanism for establishing inter-brain synchrony during mutual gaze may be mutual phase resetting (Wass et al, 2020; Leong et al., 2017). It is known that the phase of neuronal oscillations reflects the excitability of underlying neuronal populations to incoming sensory stimulation (Klimesch et al., 2007; Jensen et al., 2014). Consequently, there has been much effort expended in recent years, across a range of research fields, on exploring whether neuronal oscillations could be a key mechanism for temporal sampling of the environment (Schroeder & Lakatos, 2009; Giraud & Poepell 2012; Thut et al., 2012; VanRullen, 2016b; Ruzzoli et al., 2019). For example, some evidence suggests that sensory information arriving during high-receptivity periods is more likely to be perceived than information arriving during low-receptivity periods (Busch et al., 2009; Mathewson et al., 2009; 2010; 2011; 2012). This suggests that there is an optimal (range of) phase for perceiving information. It has been argued that if this is correct then it is logical to assume that there exist mechanisms, either endogenous or exogenous, for modulating the phase of neuronal oscillations, in order to match the temporal structure of the environmental input (van Diepen et al., 2015; Ruzzoli et al., 2019). This mechanism (phase resetting) would enable more efficient processing as information would be received during periods (phases) of high receptivity.

Empirical evidence supports the role of phase resetting as an intra-brain mechanism, facilitating neural entrainment to temporal structures within the environment, for example

speech (Giraud & Poepell 2012; Biau et al., 2015). It is therefore logical to question whether similar mechanisms also operate at the interpersonal level. For example, inter-brain synchronisation may increase within a dyad following the onset of communicative signals (such as gaze, gestures, or vocalisations) that reset the phase of both interacting partners. Here, neural oscillations in both the sender (of the social signal) and the receiver's brain that were previously random with respect to each other (low inter-brain synchrony) would be simultaneously reset in response to a communicative signal. Following this reset the neural activity of both the sender and the receiver would oscillate with more consistent variation over time (high inter-brain synchrony). We recognise that according to recent theoretical accounts that this might be classed as motor-induced neural synchrony, which occurs when the behaviour of one member of the dyad drives the neural activity of both members of the dyad (Holroyd, 2022). However, it could also be that mutual gaze onsets reset the brain activity of the sender, which precedes/ causes the behaviour of the receiver, which then causes a reset in the receiver's brain activity. Here, increases in inter-brain synchrony would be a result of both partners resetting to their own behaviours. Distinguishing between these different mechanisms is only possible using event locked analysis.

Eye movement (saccade) related potentials and ERPs

One challenge in studying the impact of mutual gaze during naturalistic social interaction on dyadic EEG is that mutual gaze onsets are time-locked to eye movements which create multiple types of artifact in the EEG (Dimigen, 2020). For example, Plöchl and colleagues (2012) showed that saccadic spike potentials (EEG potentials time-locked to small, < 1°, involuntary eye movements during fixation) typically introduce a broadband artifact in the time-frequency spectrum of the EEG, which is strongest (amplitude) in the low beta (~14-30

Hz) and gamma bands (>30 Hz) of adult EEG and typically peaks between -50ms and 150ms around the offset of a saccade. Artifact generated from eye movements can overlap in time and frequency with EEG activity presumed to be related to genuine neural activity, associated with stimulation of the retina (Gaarder et al., 1964; Billings, 1989; Dimigen, 2021). This is often referred to as the lambda response (LR) (Kazai & Yagi, 2003), which is an occipital EEG potential that can be observed when saccades are made against an illuminated contrast background (Thickbroom et al., 1991). LRs typically produce broadband time-frequency activity that is strongest (amplitude) in Alpha (8-13 Hz) and low beta (~14-30 Hz), over occipital electrodes and peaks ~100ms after the offset of the saccade (Dimigen et al., 2009; 2011). The overlapping activations introduced by eye movements can make interpretation of the data challenging, a problem which is not solved using 'standard' artifact correction procedures which fail to completely remove artifact associated with eye moments from the EEG; both in adults (Plöchl et al., 2012; Dimigen, 2020) and infants (Haresign et al., 2021). Our analyses are presented using a pipeline specially designed for the removal of eye movement artifact from naturalistic EEG data using ICA (Georgieva et al., 2020; Haresign et al., 2021; Kayhan et al., 2022). However, it is important to note that in our 2021 paper we reported that we (as arguably most of the current research in developmental neuroscience using EEG is) were unable to completely remove the activity that we assumed to be artifactually related to eye movements. Therefore, in this current work, it is likely that the sender neural responses that we are investigating are a combination of some residual artifactual activity; although as discussed above these artifacts are transient (~100ms) and therefore would only impact the initial part of the ERP waveform, and genuine neural activity; after the initial ~+150-200ms. For this reason, our primary analyses will compare sections of our data that both contain saccades, and therefore have (we assume) an identical amount of eye movement artifact in them but have different consequences (either the saccade

leads to mutual gaze, or not). Furthermore, we only compared activity in the later parts of the ERP waveform after the first 100ms.

Current study: The role of mutual gaze onsets in creating inter-brain synchrony

Our study aimed to test the hypothesis that infants are sensitive to ostensive signals during free-flowing social interactions, and that mutual gaze onsets lead to mutual phase resetting, which causes increases in inter-brain synchrony. We measured dual EEG recordings from parents and infants whilst they engaged in free-flowing social interactions and investigated intra- and inter- individual neural responses to naturally occurring moments of mutual gaze. To explore the role of turn-taking in creating inter-brain synchrony, we differentiated between two types of look onset, depending on each partners' role. 'Sender' gaze onsets were defined at a time when either the adult or the infant made a gaze shift towards their partner at a time when their partner was either already looking at them (sender mutual) or not looking at them (sender non-mutual). 'Receiver' gaze onsets were defined at a time when their partner made a gaze shift towards them at a time when either the adult or the infant was already looking at their partner (receiver mutual) or not (receiver non-mutual) (see Figure 4.1).

In order to assess how these interpersonal dynamics contributed to inter-brain synchrony we used two different measures of synchrony; First, we used a measure of concurrent synchrony, Phase locking value (PLV), that measures zero-lag, undirected synchrony. This measure would best capture changes in inter-brain synchrony that resulted from changes in both partners' brains concurrently. We also used a measure of sequential synchrony, Partially directed coherence (PDC), that measures time lagged, directed synchrony. This would best capture changes in inter-brain synchrony that resulted from changes in one partner's brain that forward predicted or lead to changes in the other partner's brain. Through this we were able to consider three main sets of research questions.

Inter-brain non-event locked analysis

For our first set of research questions, we take an analytical approach similar to Leong and colleagues' 2017 paper, in which we explore whether inter-brain synchrony is stronger overall during moments of mutual gaze. Although this was not a replication study, we attempted to translate previous structured experimental designs into a more naturalistic context. For these analyses, we preselected frequency bands and electrodes of interest, based on the findings of Leong and colleagues (2017). Consistent with these findings we expected to observe greater inter-brain synchrony, in Theta and Alpha during all moments of mutual vs non-mutual gaze. Inter-brain synchrony was measured using PLV and PDC computed over EEG data, averaged over central electrodes (C3 and C4).

Inter-brain event-locked analysis

Investigating inter-brain synchrony during social interactions as a time invariant phenomenon makes it difficult to understand the underlying mechanisms (Haresign et al., 2022). Therefore, for our second set of research questions, we wanted to explore how inter-brain synchrony changes around gaze onsets. We compared inter-brain synchrony values around sender and receiver mutual vs non-mutual gaze onsets. We expected to observe greater inter-brain synchrony (measured using PLV and PDC, in frequencies 2-18 Hz, over occipital electrodes) around mutual vs non-mutual gaze onsets. For all inter-brain analyses, we used one measure of concurrent (PLV) and one measure of sequential (PDC) synchrony in order to try to better understand how sender and receiver dynamics influence inter-brain synchrony.

Intra-brain event locked analysis

It has been suggested that one mechanism that might mediate changes in inter-brain synchrony is mutual phase resetting in response to the onset of mutual gaze (Leong et al., 2017, Wass et al., 2020). For our third set of research questions, we wanted to further investigate, mechanistically how changes in inter-brain synchrony might develop around mutual gaze onsets. To examine this, we first looked at ERPs and inter-trial coherence (ITC) around gaze onsets: comparing sender and receiver mutual and non-mutual gaze onsets. Inter-trial coherence measures consistency in phase angles over trials/ time at a single given electrode and has been used to phase resetting (Makeig et al., 2004). In comparison PLV measures the consistency in phase angle differences between two electrodes. Based on previous findings (e.g., Farroni et al., 2002) we expected to observe significant ERPs and increases in ITC relative to both sender and receiver mutual and non-mutual gaze onsets. Although during receiver non-mutual gaze onsets only, the receiver was looking at an object and not at their partner's face, research has shown that humans are highly sensitive to eye gaze in their peripheral vision (Loomis et al., 2008). ERPs and ITC were first assessed against a baseline to test whether there was a significant event-locked neural response relative to both sender and receiver mutual and non-mutual gaze onsets. We then examined whether ERPs and event locked ITC was greater around sender and receiver mutual vs non-mutual gaze onsets. We expected to observe larger ERP amplitudes, and greater ITC (in frequencies 2-18 Hz, over occipital electrodes) around mutual vs non-mutual gaze onsets.

Overall, through these analyses we wanted to explore three questions; a) do we observe above chance inter-brain synchrony during mutual gaze/ around mutual gaze onsets during free-flowing natural social interactions. B) do we observe phase resetting to mutual gaze onsets in natural contexts. C) if we observe above chance inter-brain synchrony and phase resetting around mutual gaze onsets, are they linked.

4.2. Methods

4.2.1. Ethics statement

This study was approved by the Psychology Research Ethics Committee at the University of East London. Participants were given a £50 shopping voucher for taking part in the project.

4.2.2. Participants

Of the 90 infants we tested for this study, 21 contributed no data at all, 6 contributed EEG data that was too noisy even after data cleaning and 4 were lost due to human error, e.g., failed synchronisation triggers. We also excluded all participants with fewer than 5 gaze onsets, leading to an additional 4 datasets being excluded. The final sample contained 55 healthy (23 F), *M* = 12.2-month-old (*SD* =1.47) infants, that participated in the study along with their mothers.

4.2.3. Power calculations.

For the non-event locked analysis, as this analysis were based on previous findings, we estimated the required sample size to observe a difference between the two groups (as a product of gaze type), using the G*power tool (Faul et al., 2007). For this, we used data from Leong et al., (2017) as an estimator of the expected effect size for the analysis of non-event locked synchrony ($r^2 = 0.332$). Based on an Alpha level of .05, in our sample size (*N* = 55) we had a >99% chance of observing an effect of gaze type on inter-brain synchrony of the magnitude observed in previous work.

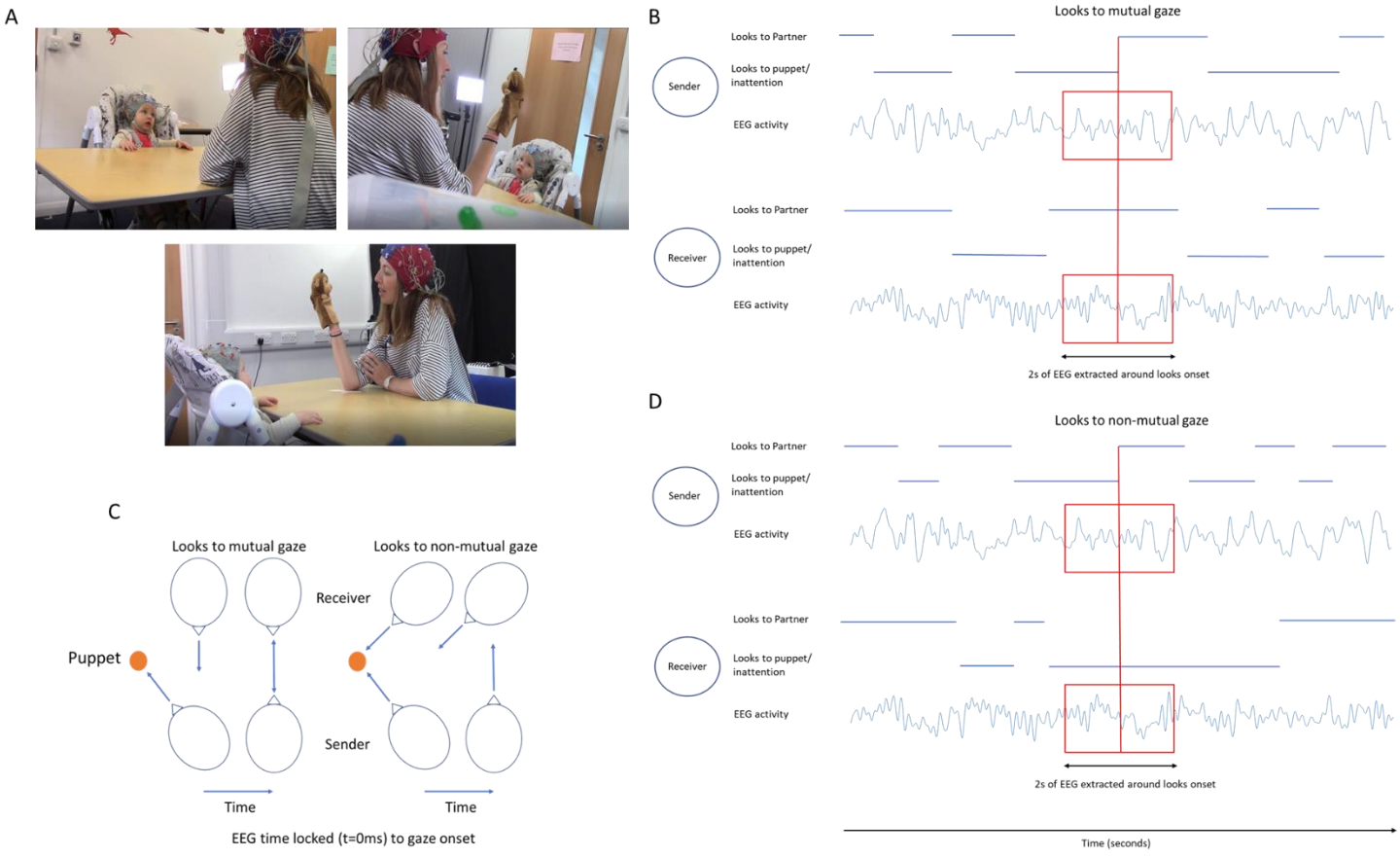


Figure 4.1. Illustration of experimental set-up and design for event locked analysis. A) shows screenshots of experimental recordings from three camera angles used. B) shows event locking process for sender and receiver EEG activity relative to mutual gaze onsets. Here receiver's gaze was maintained on the sender and the sender's gaze shifted from looking at the puppet/ inattention toward looking at the receiver, facilitating mutual gaze. +/- 1 second of EEG activity was extracted around gaze shift. C) shows topographical illustration of sender's and receiver's gaze relative to mutual and non-mutual gaze onsets. D) shows event locking process for sender and receiver EEG activity relative to mutual gaze onsets. Here receiver's gaze was maintained on either the puppet (or inattention) and the sender's gaze shifted from looking at the puppet/ inattention toward looking at the receiver, facilitating non-mutual gaze. +/- 1 second of EEG activity was extracted around gaze shift.

Infants were positioned immediately in front of a table in a highchair. Adults were positioned on the opposite side of the 65cm-wide table, facing the infant. Adults were asked to stage a 'three-way conversation' between the infant and a small hand puppet and to try to spend an equal amount of time looking at the puppet and the infant. Dual EEG was continuously acquired from the parents and infants for the approx. 5 min duration of the play session (M = 386.1, SD = 123.9 seconds)

4.2.5. Behavioural data

Video recordings were made using Canon LEGRIA HF R806 camcorders recording at 50fps positioned next to the infant and parent respectively. Video recordings of the play sessions were coded offline, frame by frame, at 50 fps. This equates to a maximum temporal accuracy of ~20ms. Coding of the infant's and adult's gaze was performed by two independent coders. Cohen's kappa between coders was >85%, which is high (McHugh, 2012). EEG was time-locked to the behavioural data offline based on the video coding using synchronized LED and TTL pulses. To verify the synchronisation, we manually identified blinks in the behavioral data and looked to see if this matched the timing of the blinks in the EEG data.

4.2.6. EEG data acquisition

EEG signals were obtained using a dual 32-channel Biosemi system (10-20 standard layout), recorded at 512 Hz with no online filtering using the Actiview software.

4.2.7. EEG artifact rejection and pre-processing

A fully automatic artifact rejection procedure was adopted, following procedures from commonly used toolboxes for EEG pre-processing in adults (Mullen, 2012; Bigdely-Shamlo, et al., 2015) and infants (Gabard-Durham et al., 2018; Debnath et al., 2020). Full details of the pre-processing procedures can be found in (Haresign et al., 2021). In brief the data was filtered between 1 and 20Hz and re-referenced to a robust average reference. Then we interpolated noisy channels based on correlation; if a channel had a lower than 0.7 correlation to its robust estimate (average of other channels) then it was removed. The mean number of channels interpolated was 3.9 (*SD* =2.1) for infants and 3.9 (*SD* =4.4) for adults. Then for the infant data only we removed sections from the continuous data in which the majority of channels contained extremely high-power values. Data was rejected in a sliding 1 second epoch and based on the percentage of channels (set here at 70% of channels) that exceeded 5 standard deviations of the mean EEG power over all channels. For example, if more than 70% of channels in each 1-sec epoch exceeded 5 times the standard deviation of the mean power for all channels then this epoch is marked for rejection. We found that for adults this step was primarily removing activity that could be removed with ICA (e.g., eye movement artifact) without removing entire sections of the data. The average amount of continuous data removed was 11.9% (*SD* =14.6%) for infants. Finally, we used ICA to remove additional artifacts.

Careful attention was paid to artifact and the amount of noise in the data throughout. In the supplementary materials (see 4.7.8) we report the results of standard measures of EEG data quality (Luck et al., 2021). Including universal measures like these enables fast and easy comparison between studies and allows the overall quality of the data to be readily assessed.

We paid particular attention to eye movement artifact. In previous work we designed a system for automatically identifying and removing artifactual ICA components in infant EEG (Haresign et al., 2021). The automated system was shown to remove most but not all eye movement related artifact time-locked to saccades. Therefore, we cannot entirely rule out that some of the activity in the sender brains relative to sender mutual and non-mutual gaze onsets is an artifact of the gaze shift.

4.2.8. Time frequency analysis- extracting power and phase

Time-frequency power and phase was extracted via complex Morlet wavelet convolution. The wavelets increased from 2 to 18 Hz in 17 linearly spaced steps and the number of cycles increased from 3-10 cycles logarithmically (this approach is generally recommended; see Cohen, 2014, chapter 13).

For all non-event locked analyses presented here, frequency bands were selected based on the bands commonly used in infant research: Theta (3-6Hz) and Alpha (6-9Hz) (Marshall et al., 2011; Leong et al., 2017; Xie & Richards, 2018; van der Velde et al., 2019; Jones et al., 2020).

4.3. Analysis

4.3.1. Inter-brain non-event locked analysis - overview

First, we wanted to explore whether inter-brain synchrony was greater during all moments of mutual vs non-mutual gaze (not relative to sender/ receiver gaze onsets). Mutual gaze was defined as times when both the adult and the infant were looking at each other. Non-mutual

gaze was defined as times when the infant was looking at the adult and the adult was not looking at the infant (or vice versa). For all non-event locked analyses, EEG was time-locked to the onset of gaze, and the length of the epoch extracted equated to the duration of the look. The average look durations were 2.7 seconds (SD = 0.9s) for mutual and 1.6 seconds (SD = 0.5s) for non-mutual gaze. All epochs were then concatenated. The mean amount of continuous data available for analysis was 66.1s (SD = 41.5s) for mutual and 28.3s (SD = 17.9s) for non-mutual gaze. Because ITC (see Cohen 2014, chapter 19) and PDC are sensitive to the amount of available data, we normalised the amount of data present per condition for each participant by identifying which gaze type had the lower number of continuous data samples (n), and re-sampling data from the other gaze type condition, taking 1:n data samples.

4.3.1.1. Inter-brain non-event locked analysis - PDC

Partial directed coherence (PDC) is based on the principles of Granger causality (Baccalá and Sameshima, 2001). It provides information of the extent to which one times series influences another. PDC is calculated from coefficients of autoregressive modelling according to:

$$PDC = \left(\frac{A_{xy}(f)}{\sqrt{a_y^*(f).a_x(f)}} \right) \quad (1),$$

where $A_{xy}(f)$ is the spectral representation of bivariate model coefficients and a_y and a_x are the spectral model coefficient from the univariate autoregressive model fit. Based on previous literature (e.g., Leong et al., 2017) we chose to compute PDC in 1-second non-overlapping sliding window. We estimated the model order for each segment using Bayesian information criteria (BIC). Model order values were then averaged for all segments. The result was a

model order of 5, the same as used by Leong and colleagues (2017), which was then used for all segments.

4.3.1.2. Inter-brain non-event locked analysis – PLV

For the non-event locked analyses, the phase locking value (PLV) was calculated within a single trial in a 1 second sliding window (e.g., Tass et al., 1998) according to:

$$PLV_t = \frac{1}{T}\left| \sum_{n=1}^{T} e^{i(\phi(t,n)-\psi(t,n))} \right| \quad (2),$$

Where T is the number of observations or time samples within the window, $\phi(t,n)$ is the phase on observation *n*, at time *t*, in channel ϕ and $\psi(t,n)$ at channel ψ.

4.3.1.3. Inter-brain non-event locked analysis – group level stats for PLV and PDC

Firstly, we assessed whether PLV and PDC values were significant compared to a baseline level. This was done by calculating the PLV/ PDC between randomly paired infant and adult dyads. We generated 1000 random infant-adult pairings in this way. We then compared the group averages of the observed PLV/ PDC values for the real infant-adult pairings against the randomly permuted distributions. Under the null hypothesis that the interbrain PLV/ PDC between infants and adults is a result of their real-time social interaction, we should observe no above chance inter-brain PLV/ PDC between randomly paired infant-adult dyads. P values were generated by first z-scoring the observed PLV values and then by evaluating the z-scored value's position on a Gaussian probability density using the Matlab function normcdf.m. For this test, we used a Bonferroni correction for multiple comparisons. This was

appropriate here as we were only testing over a limited number of predefined frequencies/ channels. Following the statistical procedure adopted by Leong and colleagues' (2017), differences in PLV and PDC between mutual and non-mutual gaze were assessed using a two-way repeated-measures ANOVA, taking gaze type and frequency as the within levels, using average, over electrodes C3 and C4, infant-to-adult PDC *(I→A)* and adult-to-infant *(A→I)* PDC values, and for Theta and Alpha bands separately. A Tukey-Kramer correction for multiple comparisons was applied.

4.3.2. Inter-brain event locked analysis

4.3.2.1. Inter-brain event locked analysis – group level stats for PLV and PDC

Firstly, we assessed whether PLV and PDC values were significant compared to a baseline level. We compared all observed time-frequency PLV/ PDC values relative to sender and receiver mutual and non-mutual gaze onsets against time-frequency PLV/PDC values time-locked to randomly inserted events within the continuous data. Differences between the real and surrogate data were assessed using cluster-based permutations statistics, using an alpha value of .05 (see SM section 4.7.5 for full details). Secondly, we examined whether PLV/ PDC values were greater for sender/ receiver mutual vs non-mutual gaze onsets. This was similarly assessed over all time-frequency points using a cluster-based permutation procedure, comparing results between different types of gaze onset. Importantly, the amount of eye movement artifact was identical between the sets of results being compared.

4.3.3. Intra-brain event locked analysis

Here we wanted to investigate whether mutual gaze onsets play a role in establishing inter-brain synchrony. We did this by investigating ERPs and ITC around sender and receiver mutual and non-mutual gaze onsets. Previous research suggests that event-locked face-sensitive neural responses are strongest over parietal/ occipital electrodes (Gao et al., 2019; Haresign et a., 2021). Therefore, for our event locked analysis we chose to focus on averaged data from a cluster of 5 parietal/ occipital electrodes (PO3, PO4, O1, Oz, O2). Additional topoplots are presented in SM 4.7.7, which support the choice of electrodes. The EEG signal was divided into events from -2500 to 2500ms (t=0 denotes the onset of gaze). The mean number of events extracted was 21.1 (SD =10.8) for infant sender/ adult receiver and 15.6 (SD =12.6) for adult sender/ infant receiver mutual gaze onsets and 9.9 (SD =6.2) for infant sender/adult receiver and 28 (SD =18.3) for adult sender/infant receiver non-mutual gaze onsets. We matched the number of events between gaze types for each participant using the procedure described in section 3.1, above.

4.3.3.1. Intra-brain event locked analysis - Inter trial coherence

Inter trial coherence (ITC) measures the consistency of frequency band-specific phase angles over trials, time-locked to a specific event. The phase coherence value is computed according to:

$$ITC = \frac{1}{N}\left| \sum_{k=1}^{N} e^{i\phi(t,k)} \right| \quad (3),$$

where N is the number of trials and $\phi(t,k)$ is the phase on trial k, at time t.

4.3.3.2. Intra-brain event locked analysis - ERPs

Following previous research (Conte et al., 2020) amplitudes of the P1, N290, and P400 ERPs were measured by calculating the change in amplitude between the peak of the component of interest and the peak of the preceding component. Also following previous recent research (Conte et a., 2020; Guy et al., 2016; 2018; Xie and Richards, 2016; 2017) we used semiautomated and individualised time window selection (Guy et al., 2021). Differences in peak amplitude were quantified using the adaptive mean approach. This process involves first identifying the peak latency of the ERP potential on a participant-by-participant basis using a broad time window. Once the peak latency has been identified an average of the activity in a 20ms window around the peak is then taken (e.g., as described in Luck, 2014). We focused on three major components relevant for face/gaze processing: the P1 component, the N170 (commonly N290 in infant EEG; Conte et al., 2020) and the P300 (commonly P400 in infant EEG; Conte et al., 2020). For the P1 component we used a time window of 0 to 200ms for adults and 100 to 300ms for infants. For the N170/ N290 component we used a time window of 100 to 300ms for adults and 200 to 400ms for infants. For the P300/P400 component we used a time window of 200 to 500ms for adults and 300 to 600ms for infants. These were selected based on visual inspection of the averaged waveforms. All ERP data were baseline corrected using data from the time window -1000 to -700ms pre-gaze onset.

4.3.3.3 Intra-brain event locked analysis – group level stats for ERPs

To test whether the onset of gaze led to significant changes in amplitude relative to both sender and receiver mutual and non-mutual gaze onsets we again used nonparametric permutation testing. Here the null hypothesis was that the timing of the gaze onset (e.g., time 0) is unrelated to the observed neural response within the time window examined. To test

this, we randomly permuted the time points of the ERPs and took the average (separately around the maximum and minimum points) of the permuted ERP in the time window 0 to +500ms. This procedure was then repeated 1000 times, randomising and reshuffling the ERP on each permutation. Finally, an estimate of the permutation p-value was calculated using the z-scoring procedure outlined in section 4.2.12. Here we used cluster-based permutation statistics, using an Alpha value of .5 to correct for multiple comparisons (time points). The results are reported in section 4.4.3.

4.3.3.4 Intra-brain event locked analysis – group level stats for ITC

Firstly, we assessed whether ITC values were significant compared to a baseline level. We compared the observed time-frequency ITC values relative to sender and receiver mutual and non-mutual gaze onsets against time-frequency ITC values time-locked to randomly inserted events within the continuous data. The number of random events was matched based on the number of real gaze onsets available for each participant. Differences between the real and surrogate data were assessed using cluster-based permutation statistics, using an Alpha value of .05. We then compared differences between sender and receiver mutual vs non-mutual gaze in ITC using the same cluster-based permutation procedure see section 4.4.4.

4.4. Results

Before turning to our main research questions, we first calculated descriptive statistics to show how gaze onsets were distributed in our sample (Figure 4.2). These results indicate that infants spent, in total, 34%/31%/35% of the total interaction looking to partner/puppet/inattentive (Fig 2A). Parents spent 62%/31%/7% of the total interaction

looking to partner/puppet/inattentive (Fig 4.2D). Overall, mutual gaze periods were longer than non-mutual gaze periods (Fig 4.2C). Differences in the frequency of gaze onsets to different types of gaze were normalised using the procedures described above.

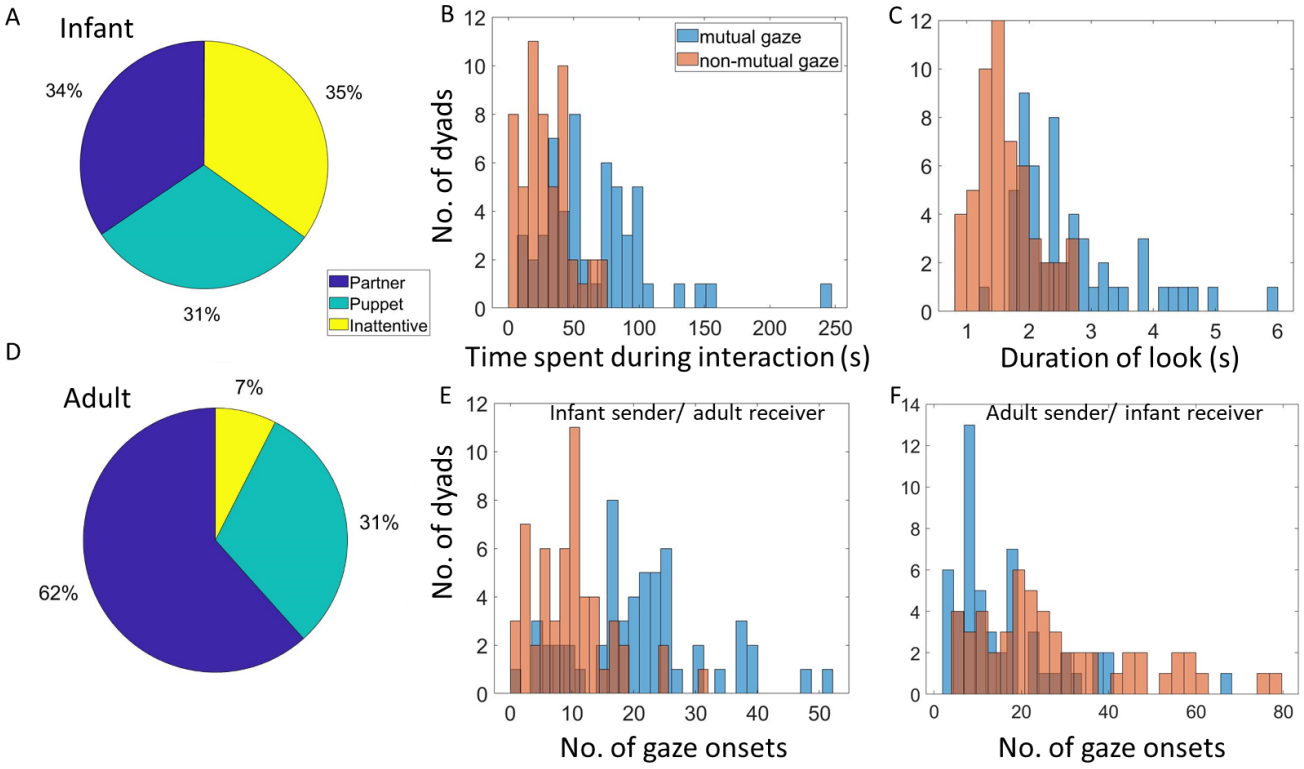


Figure 4.2. Distribution of gaze onsets in our sample. A) Distribution of time infants spent looking at different areas. B) Distribution of time dyads spend in mutual and non-mutual gaze during interaction. C) Distribution of look durations for mutual and non-mutual gaze defined from infants' look behaviour. D) Distribution of time adults spent looking at different areas. E) Distribution of infant sender/ adult receiver mutual and non-mutual gaze onsets. F) Distribution of adult sender/ infant receiver mutual and non-mutual gaze onsets.

4.4.1. Inter-brain non-event locked analysis – PLV and PDC

To investigate the relationship between inter-brain synchrony and mutual gaze, we first computed the mean PLV and PDC values across all mutual and non-mutual gaze periods in Theta and Alpha bands separately. We looked at whether PLV and PDC values were significantly greater than baseline, and then compared these values for mutual vs non-mutual gaze. Based on our previous research (Leong et al., 2017) we focused on activity over vertex electrodes (C3 and C4).

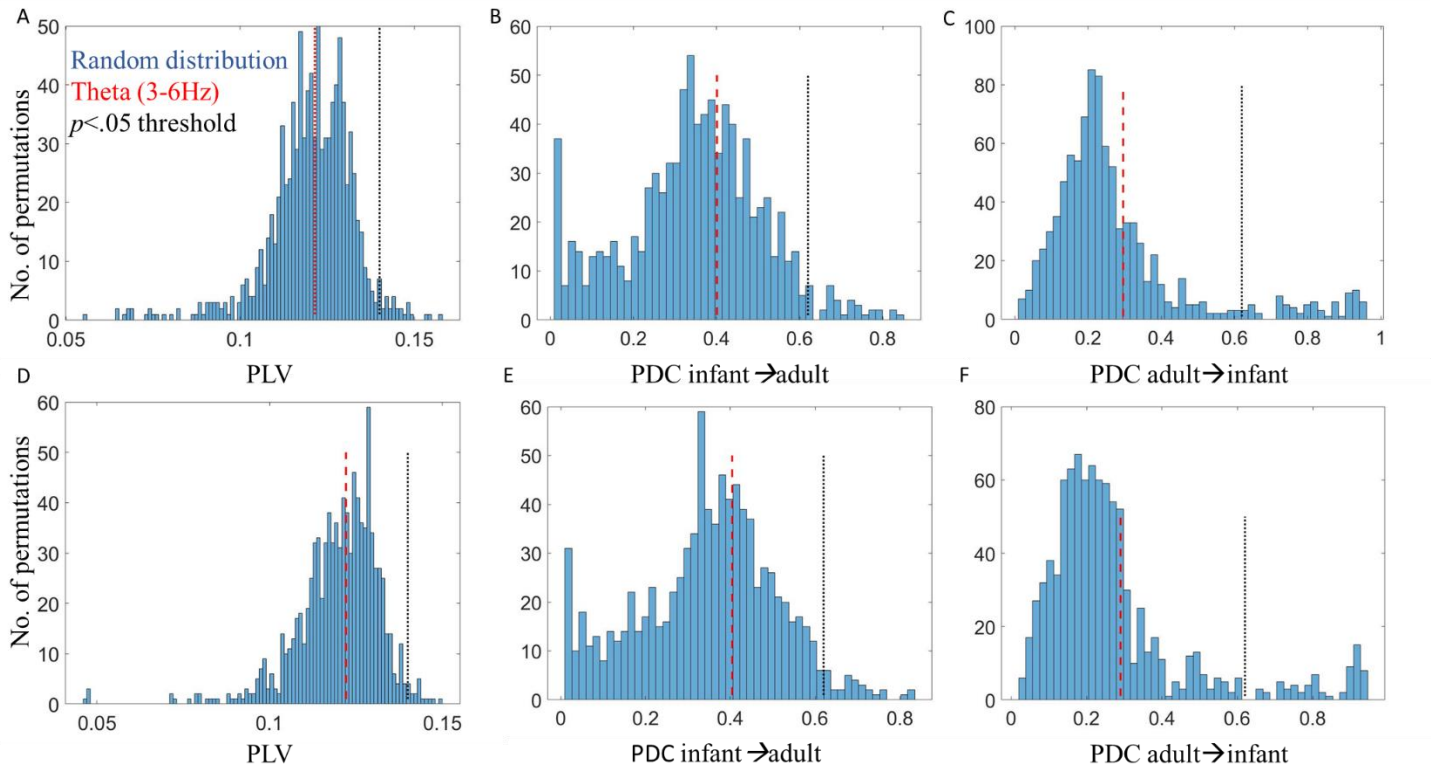


Figure 4.3. Results of permutation tests for non-event locked inter-brain synchrony analysis. A) Distribution of random pair infant-adult PLV values compared to real pair infant-adult PLV values in Theta for mutual gaze. B) Distribution of random pair infant→adult PDC values compared to real pair infant→adult PDC values in Theta for mutual gaze. C) Distribution of random pair adult→infant PDC values compared to real pair adult→infant PDC values in Theta for mutual gaze. D) Distribution of random pair infant-adult PLV

values compared to real pair infant-adult PLV values in Theta for non-mutual gaze. E) Distribution of random pair infant→adult PDC values compared to real pair infant→adult PDC values in Theta for non-mutual gaze. F) Distribution of random pair infant→adult PDC values compared to real pair adult→infant PDC values in Theta for non-mutual gaze. Red dashed lines indicate averaged observed values in Theta, black dotted lines indicate the threshold for $p < 0.05$.

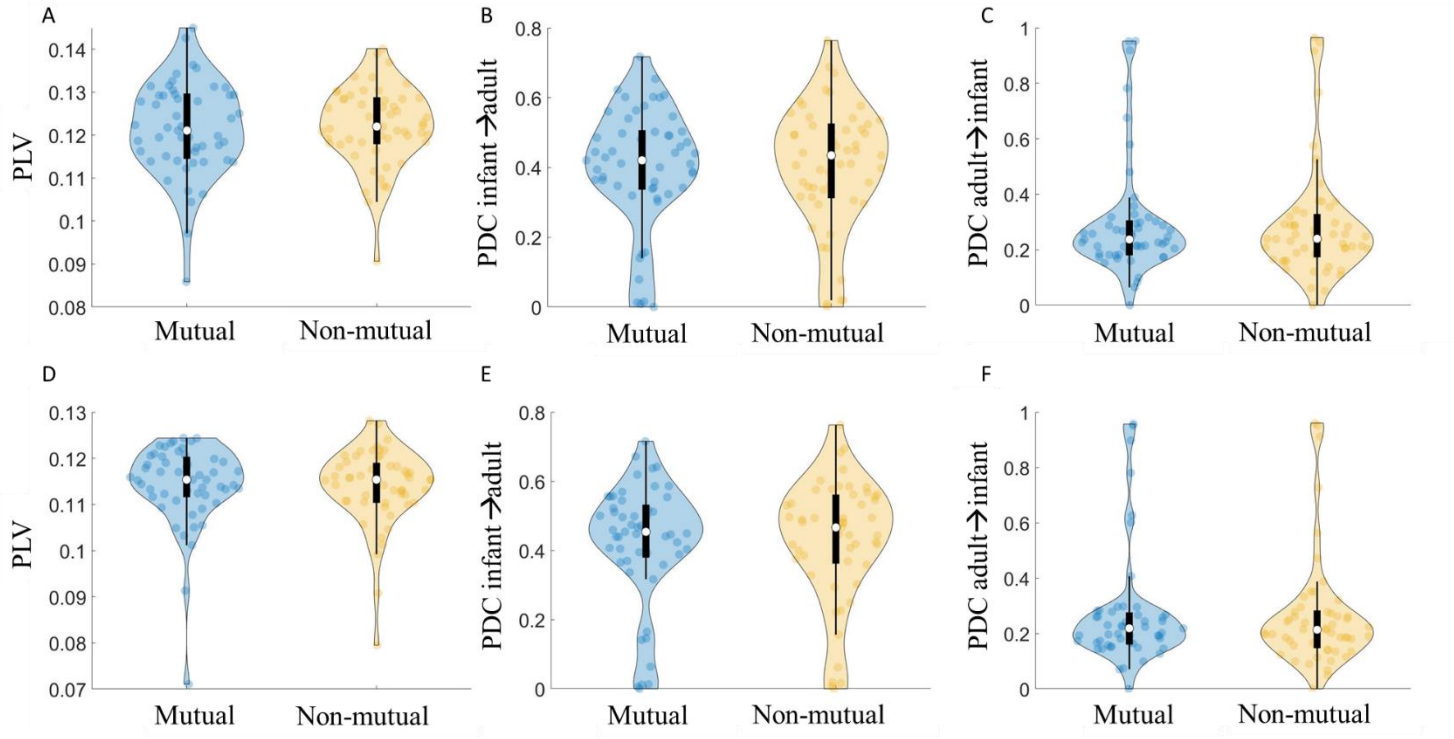


Figure 4.4. Results of non-event locked inter-brain synchrony analysis. A) Infant-adult PLV during mutual / non-mutual gaze in Theta. B) Infant→adult PDC during mutual / non-mutual gaze in Theta. C) Adult→infant PDC during mutual / non-mutual gaze in Theta. D) Infant-adult PLV during mutual / non-mutual gaze in Alpha. E) Infant-adult PLV during mutual / non-mutual gaze in Alpha. F) Adult→infant PDC during mutual / non-mutual gaze in Alpha.

Violin plots show the distribution of the data with an inset boxplot. Each point corresponds to the average PLV/PDC value of a dyad.

Figures 4.3 and 4.4 show the results of the inter-brain non-event locked analysis. We first tested whether PLV and PDC values significantly exceeded baseline values. The results of the permutation analysis indicated that PLV values did not significantly exceed baseline values: in theta (3-6 Hz) for mutual (p = 0.53) or non-mutual gaze (p = 0.56). Further, infant-to-adult PDC *(Infant→Adult)* and adult-to-infant *(Adult→Infant)* did not significantly exceed baseline values for mutual (p = 0.63/ p = 0.55) or non-mutual gaze (p = 0.64/ p = 0.54); or in alpha (6-9 Hz) for mutual (p = 0.57) or non-mutual gaze (p = 0.59). Further, infant-to-adult PDC *(Infant→Adult)* and adult-to-infant *(Adult→Infant)* did not significantly exceed baseline values for mutual (p = 0.59/ p = 0.53) or non-mutual gaze (p = 0.61/ p = 0.52).

We then wanted to test whether PLV and PDC values were greater for mutual than non-mutual gaze. As described in section 4.3.1.3, we then conducted repeated measures (RM) ANOVAs using average indices (over C3 and C4), taking frequency (2 levels) and gaze type (mutual vs non-mutual; 2 levels) as within-subjects factors. The results of the ANOVA indicated no statistically significant differences in PLV, $F(1, 55) = .04$, $p = .84$, or PDC; for infant to adult (*Infant→Adult*) influences, $F(1, 55) = .18$, $p = .68$ or adult to infant (*Adult→Infant*) influences, $F(1, 55) = .50$, $p = .48$, between mutual and non-mutual gaze. The results did indicate a significant effect of frequency (i.e., more synchrony in Theta than Alpha) for PLV, $F(1, 55) = 47.33$, $p < .01$ and PDC; *Infant→Adult*, $F(1, 55) = 41.2$, $p < .01$ and *Adult→Infant*, $F(1, 55) = 138.84$, $p < .01$). These results are summarised in Figure 4.4. Note these are uncorrected p values.

To further test the significance of gaze type on non-event locked synchrony (PLV and PDC) we calculated the Bayes Factor (BF) at the group level for both Theta and Alpha and (for PDC) including both directions of influence. We calculated $BF_{10} = p(D|H_1) / p(D|H_0)$, where D represents the data and H_1 and H_0 of the alternative and the null hypothesis respectively, using functionality from the bayesFactor toolbox (Krekelberg, 2022), based on the equations provided in Rouder et al., 2012. The BF10 tests for the presence of an effect. For all tests, the BF10 was between 1/3 and 1/10 and non-significant indicating moderate evidence (Lee & Wagenmakers, 2014) for the null hypothesis (that there is no difference between mutual and non-mutual gaze). We also converted these scores to the Bayes Factor for the absence of an effect (BF01), confirming that there was moderate to strong evidence that there was no difference between the groups. Results of our Bayes Factor analyses are given in full in SM section 4.7.6.

To summarise the results of the non-event locked analyses, suggest that mutual gaze does not induce inter-brain synchrony in this dataset and using this paradigm since a) the inter-brain synchrony for either gaze type isn't above shuffled baseline data and b) isn't different from a specifically selected other condition (non-mutual gaze).

4.4.2. Inter-brain event locked analysis – PLV and PDC

We next investigated whether onsets of mutual gaze led to changes in inter-brain synchrony and examined the sender-receiver dynamics that might contribute to this. To do this we conducted event-locked analyses with respect to gaze onsets. We first examined whether PLV and PDC values were significantly greater than baseline around gaze onsets, and then compared these values between mutual vs non-mutual gaze onsets.

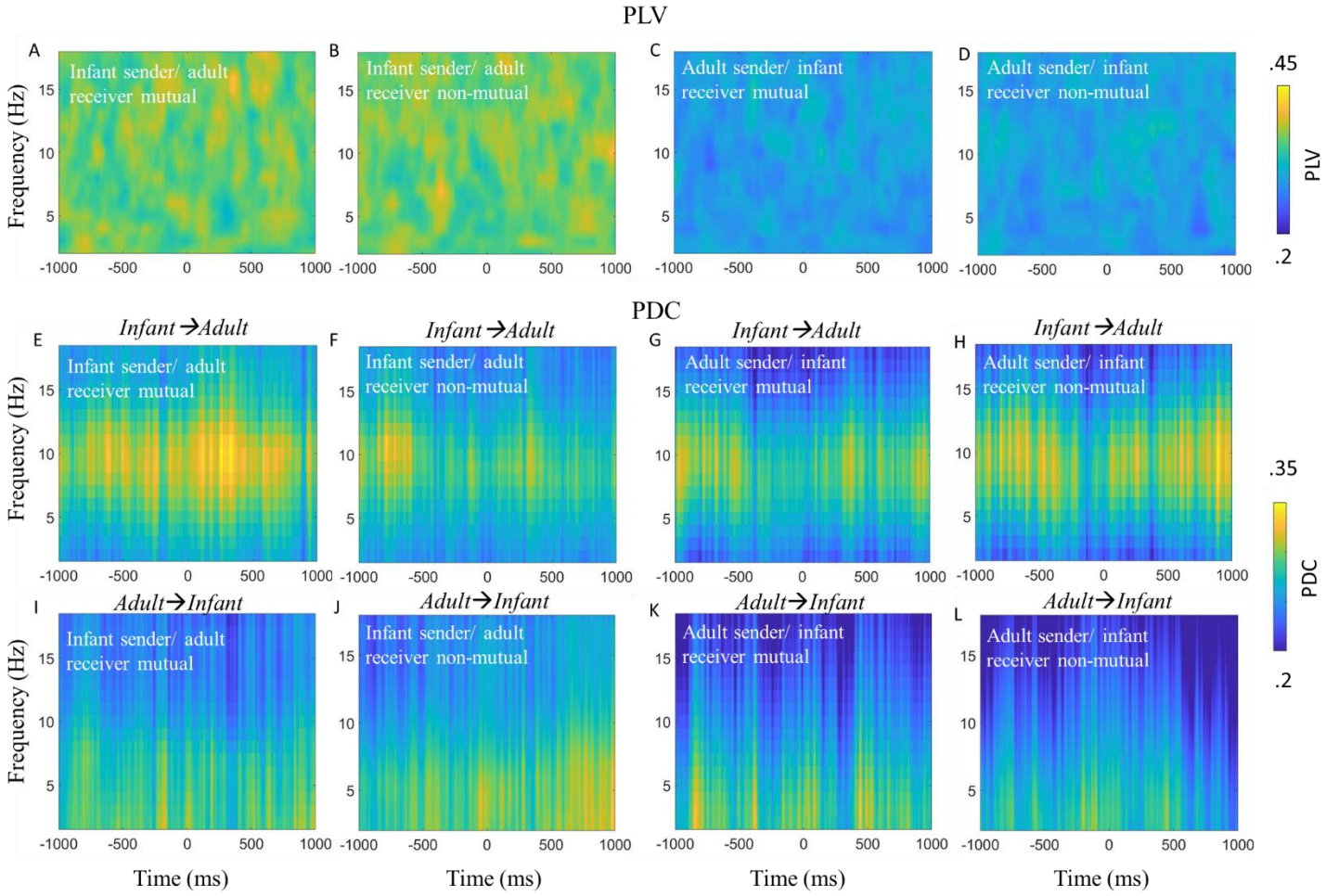


Figure 4.5. Infant-caregiver inter-brain synchrony time-locked to naturally occurring mutual and non-mutual gaze onsets. A) PLV relative to infant sender/adult receiver mutual gaze onsets. B) PLV relative to onsets of infant sender/adult receiver looks to non-mutual gaze. C) PLV relative to adult sender/infant receiver mutual gaze onsets. D) PLV relative to adult sender/infant receiver non-mutual gaze onsets. E) Infant→Adult PDC relative to infant sender/adult receiver mutual gaze onsets. F) Infant→Adult PDC relative to infant sender/adult receiver non-mutual gaze onsets. G) Infant→Adult PDC relative to adult sender/infant receiver mutual gaze onsets. H) Infant→Adult PDC relative to adult sender/infant receiver non-mutual gaze onsets. I) Adult→Infant PDC relative to infant sender/adult receiver mutual gaze onsets. J) Adult→Infant PDC relative to infant

sender/adult receiver non-mutual gaze onsets. K) Adult→Infant PDC relative to adult sender/infant receiver mutual gaze onsets. L) Adult→Infant PDC relative to adult sender/infant receiver non-mutual gaze onsets. No significant differences were identified.

We first tested whether PLV and PDC values over occipital electrodes and in the 2-18 Hz range significantly exceeded baseline values generated from a permutation procedure for infant sender/ adult receiver and adult sender/ infant receiver looks to mutual and non-mutual gaze. The result of the permutation analysis indicated that event locked PLV and PDC values around mutual and non-mutual gaze onsets were not significantly different from baseline values (see SM section 5 for full details). Therefore, we failed to reject the null hypothesis that there are no changes in PLV and PDC that are time-locked to gaze onsets.
We also observed no statistically significant differences for the effect of gaze type (e.g., mutual vs non-mutual). Therefore, we failed to reject the null hypothesis that there was no difference in PLV and PDC between looks to mutual or looks to non-mutual gaze (see Figure 4.4, and SM section 4.5).

To further test the significance of gaze type on event locked inter-brain synchrony (PLV and PDC) we calculated the Bayes Factor at the group level for both Theta and Alpha and (for PDC) including both directions of influence. For all tests, the BF10 was between 1/3 and 1/10 and non-significant indicating moderate evidence (Lee and Wagenmakers 2014) for the null hypothesis (that there is no difference between mutual and non-mutual gaze). We also converted these scores to the Bayes Factor for the absence of an effect (BF01), confirming that there was moderate to strong evidence that there was no difference between the groups. Results of our Bayes Factor analyses are given in full in SM section 4.7.

4.4.3. Intra-brain analysis - ERPs

Earlier we discussed some of the potential mechanisms that could lead to changes in inter-brain synchrony. Here, we wanted to examine intra-brain sender/ receiver dynamics around mutual gaze. To do this we compared ERPs between sender and receiver mutual vs non-mutual gaze onsets.

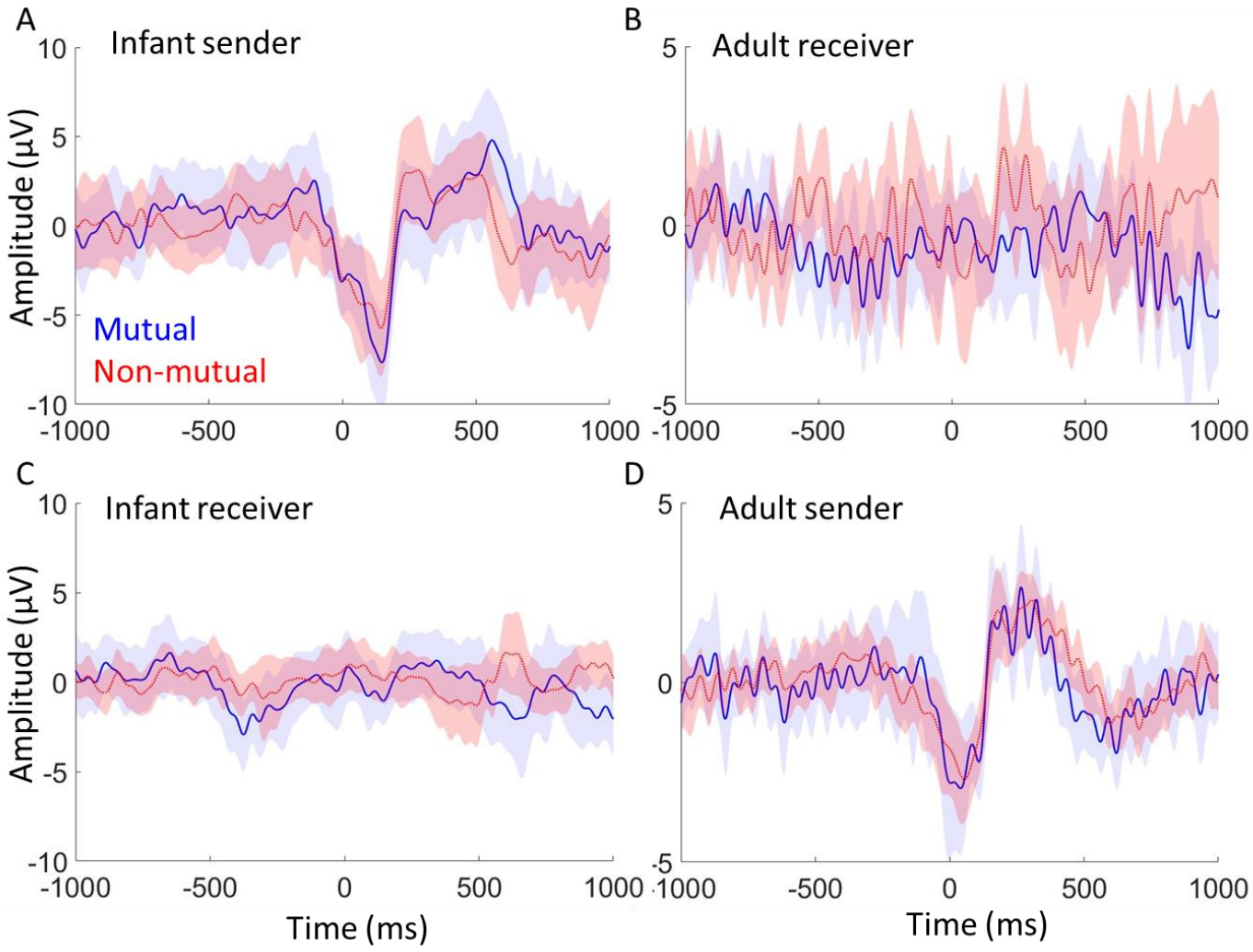


Figure 4.6. Event-related potentials time-locked to naturally occurring mutual and non-mutual gaze onsets. A) Infant ERP relative to onsets of infant sender mutual gaze and non-mutual gaze onsets. B) Adult ERP relative to adult receiver mutual and non-mutual gaze

onsets. C) Infant ERP relative to onsets of infant receiver mutual and non-mutual gaze onsets. D) Adult ERP relative to adult sender mutual and non-mutual gaze onsets. For each the shaded area indicates 95% confidence intervals; thicker lines indicate grand average waveforms. Additional topoplots can be found in SM 4.7.7.

Figure 4.6 shows the results of the intra-brain ERP analysis, comparing sender and receiver mutual and non-mutual gaze onsets. We first tested whether ERP values for sender and receiver mutual and non-mutual gaze onsets significantly exceeded baseline values generated from a permutation procedure (see section 4.3.3.3). The permutation analysis indicated that ERP amplitudes (this is just looking at whether there is a positive peak in the 0-500ms time window) in the post gaze onset window were significantly higher than baseline for sender mutual ($p < 0.01$ for both infants and adults) and non-mutual ($p < 0.01$ for both infants and adults) gaze onsets, in both infants and parents, but not for receiver mutual ($p = P_z,$ for infants and $p = 0.2$, for adults) or non-mutual ($p = P_z,$ for infants and $p = 0.6$, for adults) gaze onsets in either parents or infants.

We then compared ERP amplitudes between mutual and non-mutual gaze onsets. As the ERP amplitudes were non-significant over baseline, relative to receiver mutual and non-mutual gaze onsets we focused our comparison on sender mutual vs non-mutual gaze onsets. The results of the paired samples t-test indicated no statistically significant differences after correcting for multiple comparisons. This was consistent for all three components; P1 (p=0.66 for the infant data and p=0.04 for the adult data; uncorrected p-values), N170/N290 (p=0.61 for infants and p=0.45 for adults) and P300/P400 (p=0.21 for infants, p=0.59 in adults).

Throughout the event locked analyses careful attention was paid to what activity reflected genuine neural responses and what was related to artifact. To investigate this in more detail we performed additional analyses (see SM section 4.7.2) in which we compared the senders' neural responses pre and post artifact cleaning and compared activity over frontal vs occipital electrodes. The results of this analysis suggested that it is unlikely that these findings are driven by the eye movement artifact itself, but rather the resulting neural response. In order to further test the sensitivity of our paradigm, we also compared sender neural responses between looks to object vs looks to partner gaze (see SM section 4.7.1). The results of this analysis suggested that our paradigm differentiates neural responses to face vs object looks, consistent with the results from previous ERP studies.

4.4.4. Intra-brain event locked analysis – ITC

Lastly, we examined the possibility that the onset of mutual gaze could act as synchronisation triggers to concomitantly reset the phase of the sender and receiver's ongoing neural oscillations.

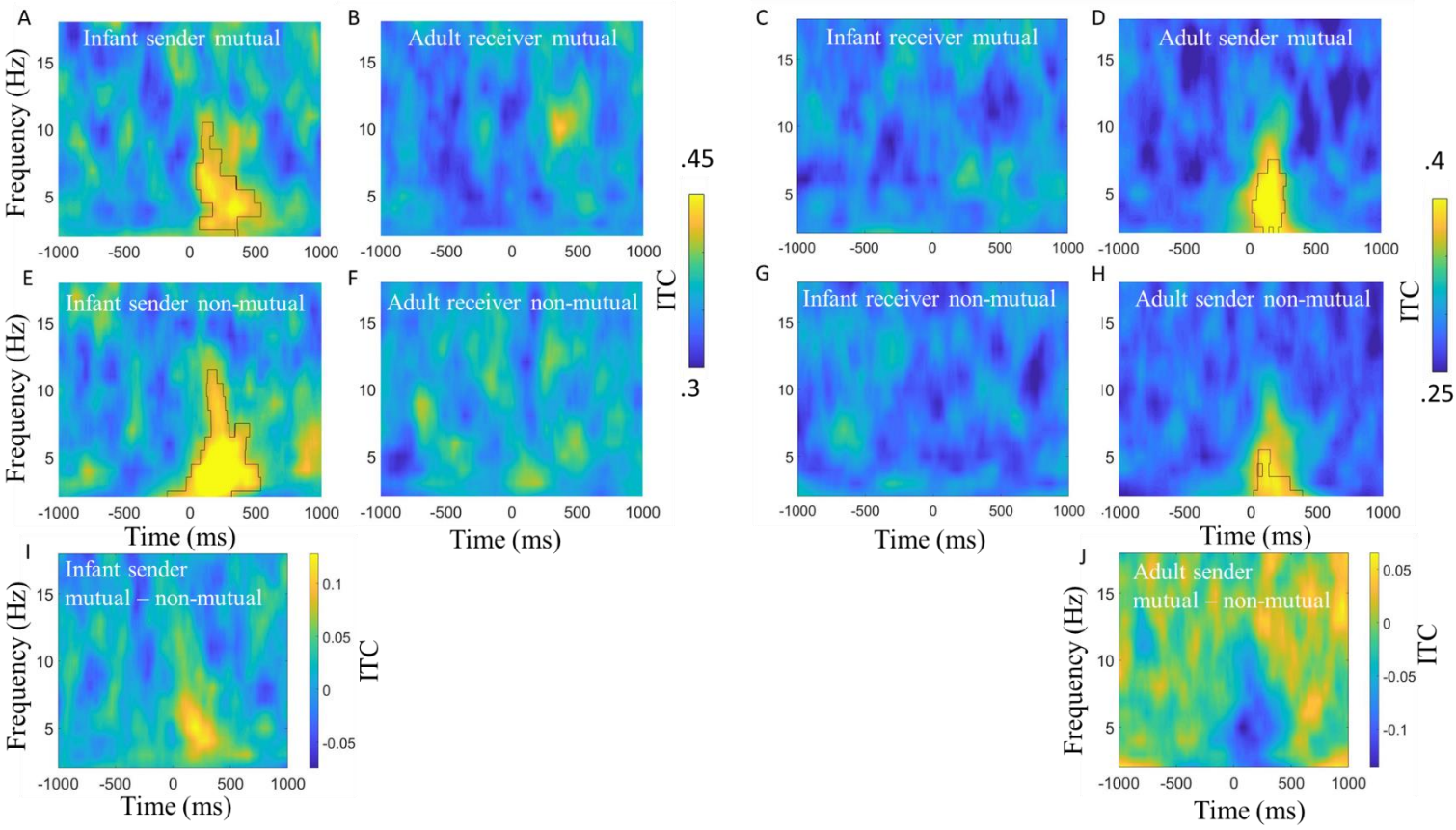


Figure 4.7. Inter-trial phase coherence time-locked to naturally occurring mutual and non-mutual gaze onsets. A) Infant ITC relative to infant sender mutual gaze onsets. B) Adult ITC relative to adult receiver mutual gaze onsets. C) Infant ITC relative to infant receiver mutual gaze onsets. D) Adult ITC relative to adult sender mutual gaze onsets. E) Infant ITC relative to infant sender non-mutual gaze onsets. F) Adult ITC relative to adult receiver non-mutual gaze onsets. G) Infant ITC relative to infant receiver non-mutual gaze onsets. H) Adult ITC relative to adult sender non-mutual gaze onsets. For all, black borders highlight activity that was significantly greater than baseline after cluster correction for multiple comparisons

using p = .05. I) Difference plot between infant sender mutual vs non-mutual gaze onsets. J) Difference plot between adult sender mutual vs non-mutual gaze onsets. Hotter colours indicate more ITC for mutual vs non-mutual.

Figures 7 shows the results of the event-locked ITC analysis, comparing sender and receiver mutual and non-mutual gaze onsets for infant and adults separately. We first tested whether ITC values, over occipital electrodes, and frequencies 2-18 Hz, significantly exceeded baseline values generated from a permutation. The permutation analysis indicated that ITC in the post-gaze onset window was significantly higher than baseline for sender mutual (Figure 4.4 A and D) and non-mutual (Figure 4.4 E and H) gaze onsets, in both parents and infants, but not for receiver mutual (Figure 4.4 B and C) and non-mutual (Figure 4.4 F and G) gaze onsets in either parents or infants.

We then wanted to test whether ITC values were greater for mutual vs non-mutual gaze onsets. The results of the cluster-based permutation analysis indicated no statistically significant differences between looks to mutual vs non-mutual gaze in the sender's brain activity in either parents or infants. The permutation analysis did reveal that ITC in the post gaze onset window was significantly greater relative to infant, but not adult receiver mutual vs non-mutual gaze onsets (peaking in Theta between 0-500ms post gaze onset, full details can be found in SM section 4.7.4).

4.5. Discussion

We took dual EEG recordings from parents and infants whilst they engaged in naturalistic free-flowing social interactions. Our data were analysed using cleaning and analysis procedures specially designed for naturalistic dual EEG data (Haresign et al., 2021; 2022). Since our analyses suggested that eye movement artifact cannot be completely removed from the EEG, we primarily compared sections of the data that were both identically time-locked to saccades, and therefore contain (presumably) an identical amount of eye movement artifact. The saccades were differentiated by the consequences of the saccade (either the saccade leads to mutual gaze, or not). Furthermore, we also compared activity only in the later parts of the ERP waveform after the first 100ms, when we were confident that no residual artifact remained.

Overall, through our analyses we wanted to explore three questions; a) do we get above chance inter-brain synchrony during mutual gaze/ around mutual gaze onsets during free-flowing natural social interactions. B) do we get phase resetting to mutual gaze onsets in natural contexts. C) if we get above chance inter-brain synchrony and phase resetting around mutual gaze onsets, are they linked. As the result of our analysis indicated that inter-brain synchrony was not greater than chance during mutual gaze/ around mutual gaze onsets during natural parent-infant social interactions we were not able to answer question c. The discussion focuses on the analysis conducted to try to address these questions.

Mutual gaze and inter-brain synchrony

Our first set of research questions explored whether inter-brain synchrony was greater during mutual vs non-mutual gaze. Our results indicated that inter-brain synchrony did not significantly exceed baseline values for either mutual or non-mutual gaze. Further, comparing mutual vs non-mutual gaze, our results indicated that inter-brain synchrony was not greater during mutual vs non-mutual gaze, contrary to what we had hypothesised. These null findings were consistent across both frequencies and both measures of synchrony that we looked at (PLV and PDC), and across both our non-event and event locked analyses.

These results are inconsistent with previous studies that observed greater inter-brain synchrony during continuous (i.e., not relative to specific behaviours/ events within the interaction, but rather looking across all moments of a given behaviour during social interaction) moments of mutual vs non-mutual gaze. For example, in our previous paper we found increased inter-brain synchrony using PDC, in Theta and Alpha, over C3 and C4 electrodes in $N = 29$ *8*-month-olds (Leong et al., 2017). In the present study, we measured PDC and PLV across the same frequencies and electrodes in $N = 55$ *12*-month-olds.

Although we followed the same analytical techniques as Leong et al., 2017, we used different pre-processing techniques and a different (less structured, more naturalistic) paradigm, which could explain why our results differ. Firstly, the previous study featured an unfamiliar live adult singing nursery rhymes to an infant. Our present study, in contrast, featured primary caregivers interacting freely with the infant, using a puppet that they held in their hand. Infants' sensitivity to novel interaction partners is well documented (Bushnell et al., 1989; de Haan & Nelson, 1997; 1999; Barry-Anwar et al., 2016; Hoehl et al., 2012). Therefore, one explanation for the positive effects of gaze type on inter-brain synchrony in our previous study could be due to the saliency of mutual gaze in the presence of an unfamiliar adult. To

investigate this further we performed the same analysis with data collected from infant-adult dyads. Here the infants interacted with an unfamiliar adult (one of two research assistants). The results of these analyses are reported in full in SM 9, but summarised here, we found consistent with our main analyses of infant-caregiver dyads inter-brain synchrony was not above chance around mutual gaze onsets and did not differ between mutual and non-mutual gaze onsets. Further we found that phase resetting around mutual gaze onsets was strongest for infant and adult sender compared with receiver gaze onsets.

Second, in our previous study, adults continuously sung nursery rhymes to the infants during the interactions, whereas in our present study they talked normally. As sung nursery rhymes are highly periodic (Suppanen et al., 2019) and evidence suggests that infant's neural activity entrains to the temporal structure of these songs (Leong et al., 2017a; Attaheri et al., 2022), it could be that the regularity of the nursery rhymes introduced an external periodic stimulus into the environment that was driving the inter-brain entrainment (e.g., Perez et al., 2017). Here, mutual gaze might only enhance or maintain synchrony that is already established, by facilitating shared attention and therefore upregulating attention-enhanced neural synchrony. It will be important for future research to examine inter-brain synchrony in a variety of settings, ranging from very unstructured settings such as those used in the present study to more structured settings in which there are environmental stimuli with more regular and predictable inputs.

Additionally, to address the possibility that differences in pre-processing procedures could explain why we failed to replicate previous findings we cleaned our data following to the best of our ability the pre-processing procedures outlined in Leong and colleagues' (2017) study. The full results of this are reported in SM 4.7.10, but to summarise we found that cleaning

the data following the procedures of Leong and colleagues had no impact on the significance of any of the results of the main paper, ruling out the possibility that differences pre-processing procedures might be the cause of the discrepancy. Overall, these inconsistencies highlight the likely context-specific and localised nature of inter-brain synchrony, and further emphasise the importance of replication and standard data quality measures (Luck, 2021) when studying inter-brain dynamics (Holroyd, 2022).

Phase resetting around gaze onsets

For our second set of research questions, we explored event-locked intra and inter-brain neural responses associated with mutual gaze onsets. Through this, we aimed to test our previously published hypothesis that concomitant phase resetting in the sender and the receiver's brain at the onset of gaze may drive inter-brain synchrony (Leong et al., 2017; Wass et al., 2020). Overall, the results of our event-locked analyses are inconsistent with this idea. Contrary to our hypothesis, inter-brain synchrony did not significantly exceed baseline values for sender/ receiver mutual or non-mutual gaze onsets and was not significantly different between sender or receiver mutual vs non-mutual gaze onsets. Further, whilst we found that sender but not receiver mutual and non-mutual gaze onsets led to significant increases in ITC and amplitude (ERPs) over baseline, we did not find significant differences between sender or receiver mutual vs non-mutual gaze onsets. We did, however, find evidence for increases in ITC relative to sender mutual and non-mutual gaze onsets (section 4.3.4); but it is difficult to conclude that this represents phase resetting of brain oscillations. It could also be that changes in event locked amplitude/ power create the artifactual appearance of phase synchrony (Muthukumaraswamy et al., 2011) – a fact that the close correspondences we observed between ITC and event-locked changes in amplitude/power (see SM 4.7.2) would appear to support.

One possible driver of the sender neural responses could be residual eye movement artifact in our data. In the supplementary analysis (SM section 4.7.3) we compare time-frequency power over frontal and occipital electrodes before and after ICA cleaning and report that ICA cleaning removed most, but not all, of the assumed artifactual activity associated with the eye movement- a conclusion consistent with our previous research (Haresign et al., 2021). This analysis also allowed us to identify that these artifacts are transient (~100ms) and therefore only impacted the initial part of the ERP waveform. After the initial ~+150-200ms we observed ERP components that look very similar to ERPs observed in traditional screen-based tasks, with clear P1, N290 and P400 components. For added safety, however, our main analyses were based on comparing sections of the data that are both identically time-locked to saccades, and therefore contain an identical amount of eye movement artifact.

Overall, then, our results challenge the theory that phase resetting around key communicative signals such as mutual gaze is a mechanism through which inter-brain synchrony is achieved. Assuming that inter-brain synchrony according to more recent frameworks (Holroyd, 2022) is associated with mutual gaze. This points to the potential importance of other potential drivers of inter-brain synchrony, that future work should investigate in more detail – such as correlated changes in amplitude/ power or changes in oscillatory frequency independent of phase resetting (see Haresign et al., 2022 for a detailed discussion), and other more periodic behaviours (e.g., speech; Leong et al., 2017a; Attaheri et al., 2022).

Re-examining the importance of 'receiver' mutual gaze in infant-caregiver social interaction

Our third aim was to test the hypothesis that infants are highly sensitive to ostensive signals during free-flowing social interactions with their caregivers. A number of influential papers

(Farroni et al., 2002; 2004 Grossman et al., 2007; Senju & Johnson 2009; Csibra & Gergely, 2009) argue that, from shortly after birth, infants' brains are sensitive to receiving ostensive signals, and that 'sender' communicative signals play a key role during naturalistic learning exchanges. However, as we noted these findings have not replicated well in developmental research; for example, Elsabbagh and colleagues (2009) or in research with adults (e.g., Watanabe et al., 2001; Taylor et al., 2001b; Watanabe et al., 2002; Itier et al., 2007; Conty et al., 2007; Ponkanen et al., 2011).

Contrary to expectations we found robust neural responses only in the senders' (i.e., the agent initiating the gaze episode) and not in the receivers' neural responses. This was true both for receiver non-mutual gaze onsets (where receivers were not looking at their partners and thus may have failed to detect their partners' gaze shift), but also for the receiver-mutual condition (where receivers were directly gazing towards their partner at the time of the gaze shift). Evidence from adult ERP studies in which dynamic changes in gaze are simulated, through the presentation of a series of static images on a screen suggest that human adults are sensitive to 'dynamic' changes in gaze (Latinus et al., 2015; Stephani et al., 2020). However, these simulated changes in gaze are still far from the continuous way that gaze is processed during real social interactions, and it is likely that the effects observed in these studies are largely driven by more low-level properties of the simulation (e.g., retinal stimulation evoked by the presentation of a series of static images) rather than reflecting the actual processing of the gaze shift. When scrambled control images are presented in this same way this produces similar neural responses to those associated with processing simulated changes in gaze (Rossi et al., 2014). This suggests that whilst these studies do capture some neural mechanisms that are sensitive to moment-to-moment changes in the visual input from our environment, these studies do a poor job of simulating the continuous flow of gaze information that happens

during real life social interaction. However, these studies do show some subtle neural sensitivity to changes in gaze orientation that perhaps we were unable to capture with the level of sensitivity afforded in our current approach. This raises basic questions over where, when and under what circumstances changes in a partner's gaze during free-flowing social interactions impacts the neural activity of the receiver (the person viewing the gaze shift). Again, one possible explanation for the inconsistencies between previous screen-based tasks and the present study is simply it is just a result of increased artifact through the use of a naturalistic paradigm.

However, we note that: i) our ERPs show a close visual correspondence with ERPs observed in traditional ERP paradigms (see Figure 4.6); ii) the overall measures of EEG data quality we reported show good quality data (see SM section 4.7.8); iii) we did replicate the findings from screen-based ERP research that infants show enhanced ERPs to images of faces vs objects (Guy et al., 2016; 2018; Peykarjou & Hoehl, 2013; Xie & Richards, 2016) (SM section 4.7.1); We observed statistically greater occipital ERP amplitudes for faces vs objects for the N290 component, but not for P1 or P400 components. Overall, then, we found changes in brain activity only in the person that initiated the episode of mutual gaze (the sender), and no changes in the recipient of the mutual gaze. This conclusion suggests a different account of the dyadic mechanisms involved in the processing of mutual gaze. In contrast to evidence suggesting that mutual gaze involves both infants and adults reciprocally influencing each other's neural activity towards shared rhythms (Leong et al., 2017), we found changes in brain activity only in the person that initiated the episode of mutual gaze (the sender). This suggests that in this context, mutual gaze processing is more supported through more basic changes at the intra-brain level, which do not specifically affect inter-brain associations. For example, evidence suggests that eye movements lead to low frequency

phase reorganisation in brain structures such as the hippocampus that are deeper than those that can be measured using scalp EEG (e.g., Hoffman et al., 2013). Eye movements may create transient increases in neural sensitivity within certain structures within an individual's brain that support them in processing the new visual information (e.g., mutual gaze) (Klimesch et al., 2007).

4.6. Conclusion

We investigated the possibility that concomitant phase resetting in response to mutual gaze onsets during naturalistic infant-caregiver interactions might be a mechanism through which inter-brain synchrony is established. We found no evidence for changes in inter-brain synchrony around gaze onsets and no evidence to support our previously published suggestion that phase resetting in the sender and the receiver's brain around mutual gaze onsets may be a mechanism through which inter-brain synchrony arises (Leong et al., 2017; Wass et al., 2020). Further, contrary to our prediction, we found that mutual gaze onsets associated with neural responses in 'senders', but not in 'receivers' brains. Overall, our study challenges current views on the importance of mutual gaze. It highlights the fact that we need pluralistic approaches to better understand early social cognition. And it highlights the importance of studying how infants perceive communicative signals during naturalistic interactions, and across different real-world contexts.

4.7. Supporting materials for chapter 4

Contents

4.7.1. Intra brain analysis – ERPs to faces vs objects

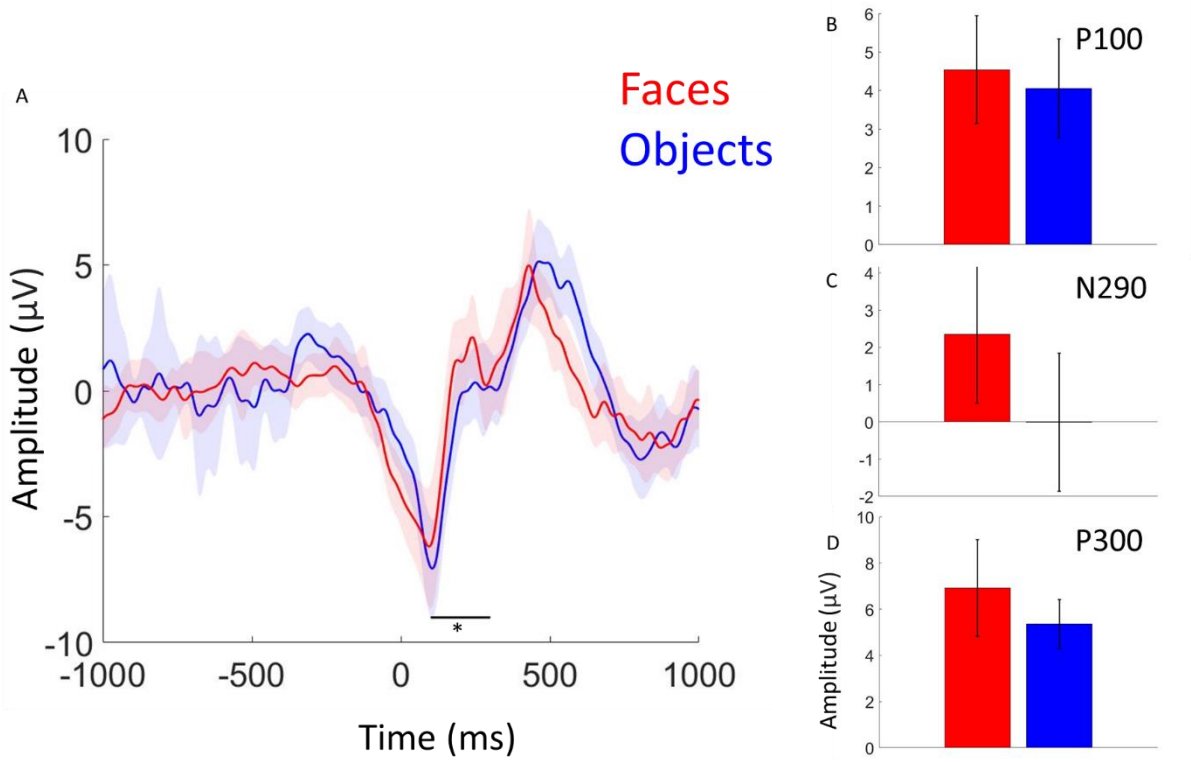


Figure 4.8. Event-related potentials time-locked to naturally occurring partner (face) gaze and object gaze onsets. A) Infant occipital ERPs relative to onsets of infant looks to their partners face vs infant looks to object. Shaded areas indicate 95% confidence intervals,

thicker lines indicate grand average waveforms and black line indicates time points at which ERP amplitudes for faces vs objects were significantly greater at p = .05.B) Grand average amplitudes for P100 component for faces vs objects. C) Grand average amplitudes for N290 component for faces vs objects. D) Grand average amplitudes for P300 component for faces vs objects. Bar charts are for visual purposes only to show the direction of the difference of the three ERP components assessed from panel A. Error bars indicate 95% confidence intervals.

Figure 4.8. shows the results of the ERP analysis, comparing infant-initiated partner and object looks. As an additional analysis to our comparison between mutual and non-mutual gaze we wanted to replicate previous findings (e.g., Conte et al., 2020) that show greater infant ERP amplitudes to faces vs objects. We observed statistically greater occipital ERP amplitudes for faces vs objects for the N290 component (p = 0.03), but not for P1 or P400 components, p = 0.14/ p = 0.50 respectively. Note these represent uncorrected p values. Our results suggested that our paradigm differentiates neural responses to face vs object looks, consistent with the results from previous ERP studies.

4.7.2. Intra brain event locked analysis - Power

It is known that changes in spectral power resulting from evoked neural responses can give the appearance of increased phase-locking/ resetting, due to changes in signal to noise ratios and error associated with estimating phase (Muthukumaraswamy et al., 2011). Power was obtained as the square of the absolute values derived from the wavelet convolution procedure described in section 4.3.4 of the main text. Time-frequency power was baseline normalised

(decibel normalised) using activity in the -1000 to -700ms time window and averaged over trials.

Differences between mutual and non-mutual gaze onsets in spectral power were assessed following the same statistical procedure as detailed in the main text, using a cluster-based correction for multiple comparisons with an alpha value of .05. We found no significant differences in spectral power between gaze types after correction for multiple comparisons (see Figure 4.10). However, it is important to note the visual similarities in the time-frequency characteristics between the observed event locked ITC (see Figure 4.9) and power changes around gaze onsets. The significance of this is discussed within section 4.5.2 of the main text.

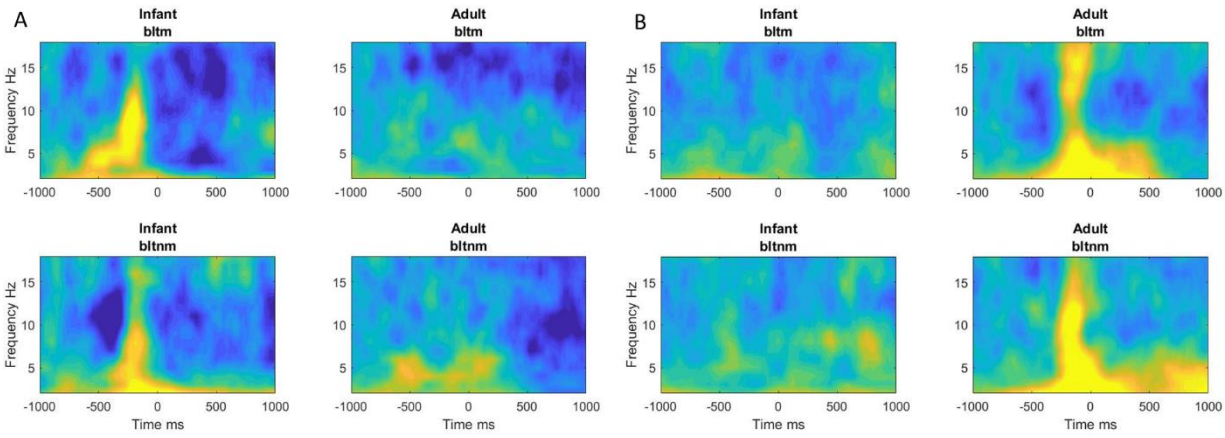


Figure 4.9. Occipital power time-locked to naturally occurring mutual and non-mutual gaze onsets. A) Infant and adult occipital power relative to infant sender/ adult receiver looks to mutual (bltm) and non-mutual gaze (bltnm) onsets. B) Infant and adult occipital power relative to adult sender/ infant looks to mutual (mltm) and non-mutual gaze (mltnm) onsets.

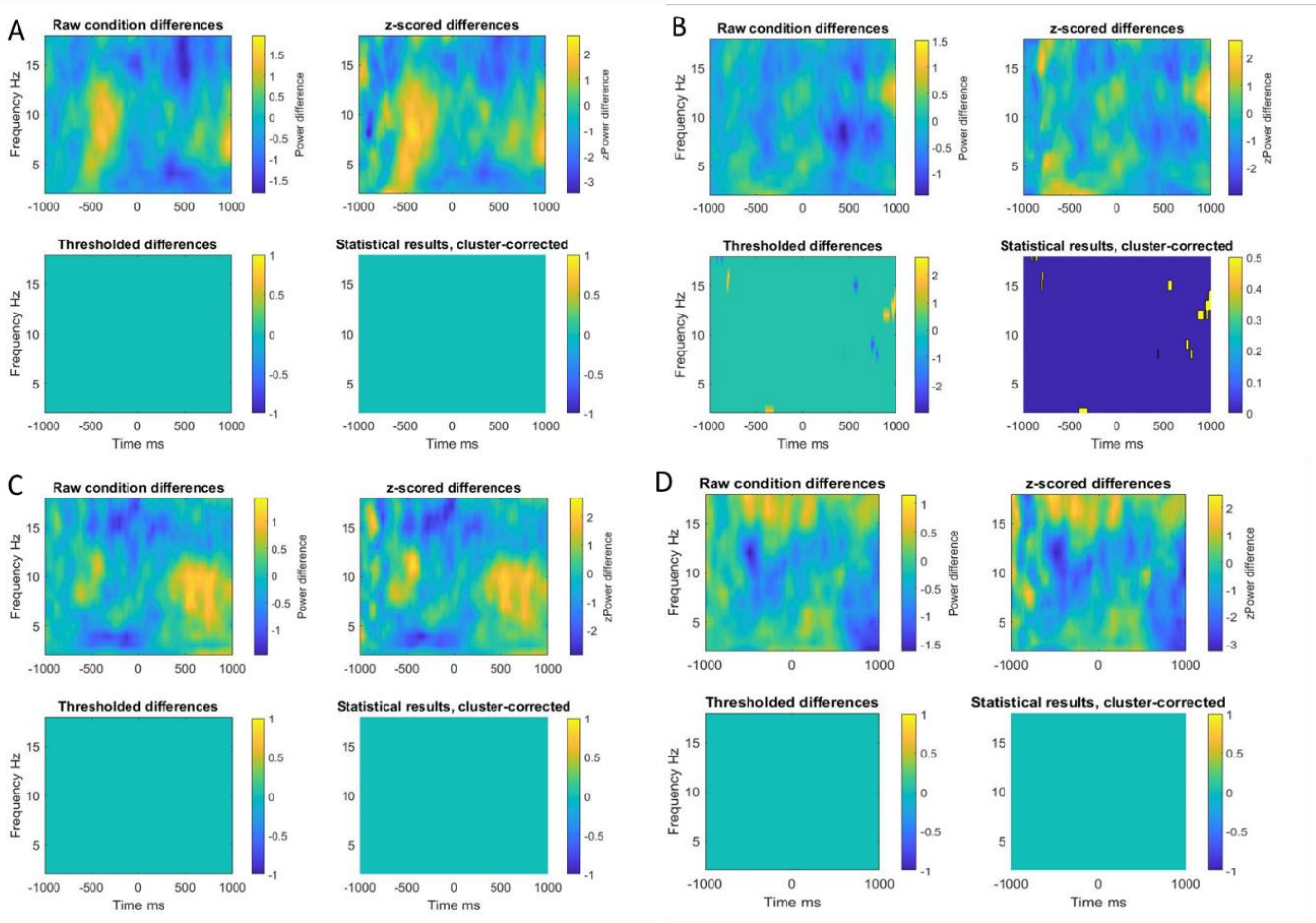


Figure 4.10. Result of cluster-based permutation procedure for event locked power for gaze type A) Result of permutation for infant occipital power comparing infant sender/adult receiver looks to mutual vs non-mutual. B) Result of permutation for adult occipital power comparing infant sender/ adult receiver mutual vs non-mutual gaze onsets. C) Result of permutation for infant occipital power comparing adult sender/ infant receiver mutual vs non-mutual gaze onsets. D) Result of permutation for adult occipital power comparing adult sender/ infant receiver mutual vs non-mutual gaze onsets. For each, the top left plot is the raw differences where hotter (more yellow) colours indicate more power for mutual vs non-mutual. The top right plot shows z scored differences. The bottom left plot shows thresholded differences (at p = .05) before cluster correction for multiple comparisons. The bottom right

shows the final results of cluster-based permutation statistics after correction for multiple comparisons.

4.7.3. ITC effects and eye movements

Here we conducted additional analysis to explore the possibility that the observed ITC effects could be a result of the eye movement artifacts and not the neural response associated with the processing of the new gaze information. To do this we first examined how the ITC effects varied topographically. Specifically, we looked at whether ITC was stronger over frontal (averaged over electrodes Fp1, Fp2, Af3, Af4) vs occipital (averaged over electrodes O1, O2, Oz, Po3, Po4) channels. From Figure S3 it can be seen that the ITC is weaker over frontal than occipital channels, suggesting that the occipital ITC is more strongly associated with the neural response rather than the eye movement artifact as we observed similar topographical patterns for amplitude and power (e.g., see SM section 4.7.2).

However, it is possible that our pre-processing procedure was bias toward removing frontal activity and therefore could have created the topographical differences we observed for ITC. In our data, we attempted to remove eye-movement artifacts using ICA decomposition, and an automated procedure was adopted for judging which components should be removed (Marriott Haresign et al., 2021). Though ICA algorithms are not biased towards removing specific types of artifacts, eye movement artifacts are often more stereotyped and easier to identify in an ICA decomposition (Chaumon et al., 2015), which could have led to biases in the systems judgement of which ICA components should be removed. Therefore, we also checked the topographical distribution of the observed ITC effects in data that had not been cleaned using ICA. Here we found consistent with our main findings, ITC was strongest over

occipital electrodes. Overall, these data suggest that it is unlikely that these observed ITC is driven by the eye movement artifact itself, but rather the resulting neural response.

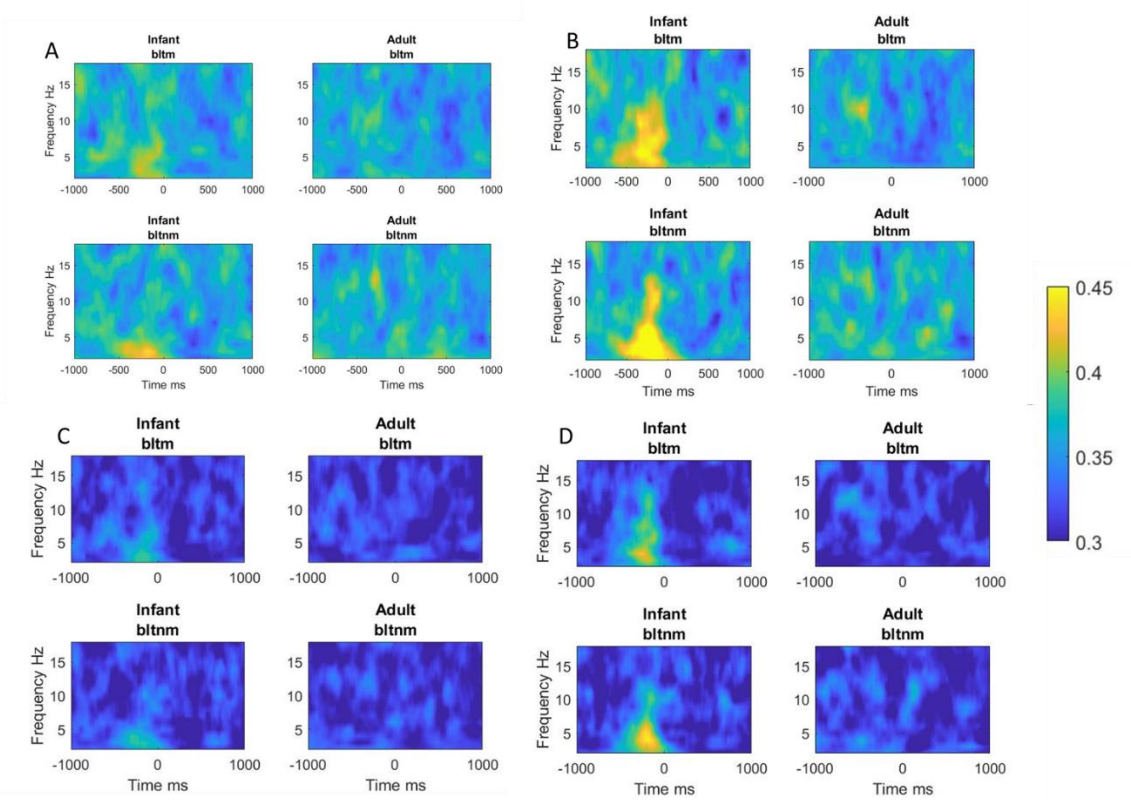


Figure 4.11. Event-locked ITC time-locked to naturally occurring mutual and non-mutual gaze onsets before and after ICA cleaning. A) Infant and adult frontal ITC relative to infant sender/ adult receiver mutual and non-mutual gaze onsets post-ICA. B) Infant and adult occipital ITC relative to infant sender/ adult receiver mutual and non-mutual gaze onsets post-ICA. C) Infant and adult frontal ITC relative to adult sender/ infant receiver mutual and non-mutual gaze onsets pre-ICA. D) Infant and adult occipital ITC relative to adult sender/ infant receiver mutual and non-mutual gaze onsets pre-ICA.

4.7.4. Intra-brain event locked analysis ITC – Results of permutation procedure

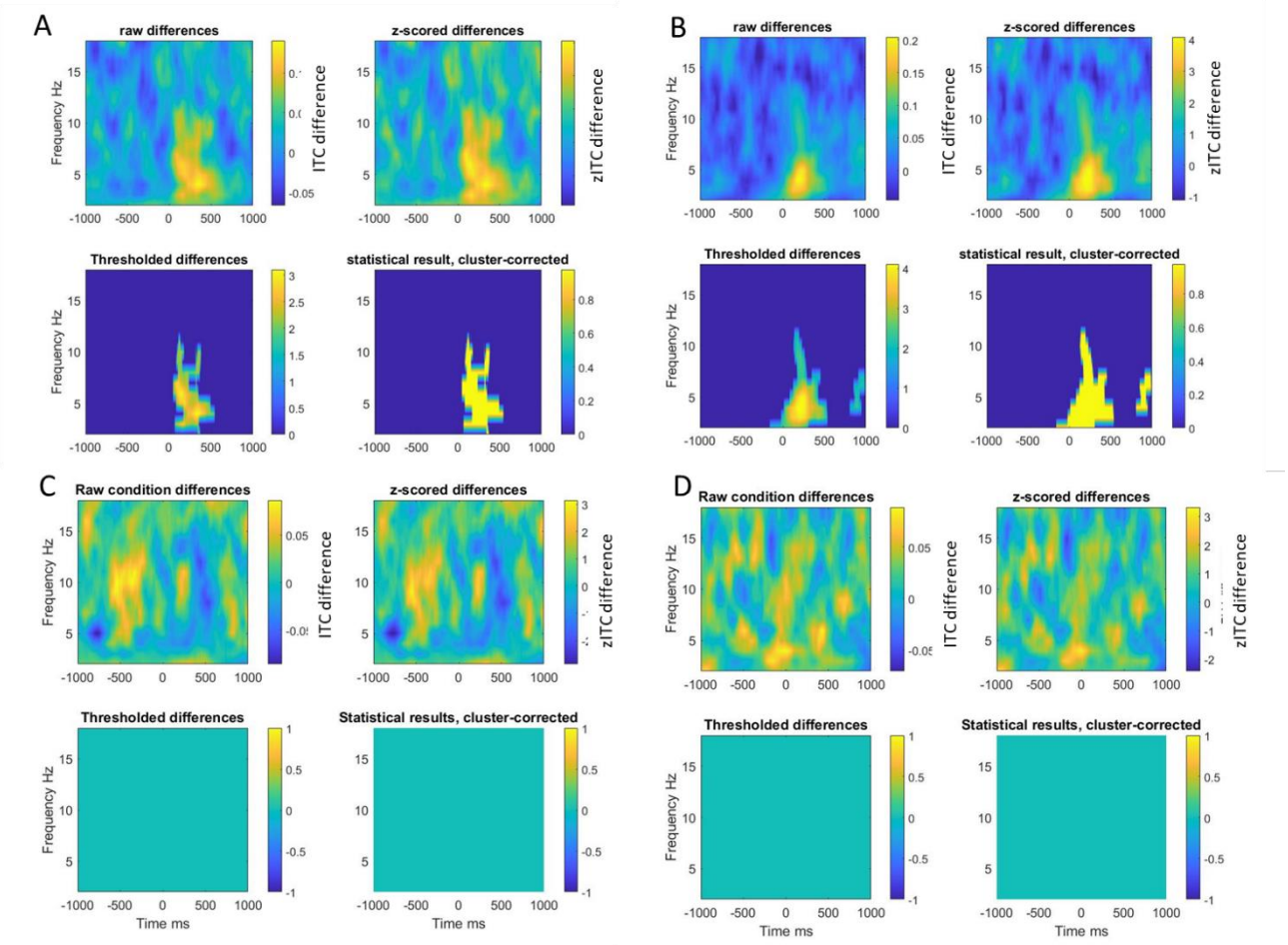


Figure 4.12. Result of cluster-based permutation procedure for event locked ITC relative to baseline. A) Result of cluster-based permutation procedure for infant occipital ITC relative to infant sender/adult receiver mutual gaze onsets. B) Result of cluster-based permutation procedure for infant occipital ITC relative to infant sender/adult receiver non-mutual gaze onsets. C) Result of cluster-based permutation procedure for infant occipital ITC relative to adult sender/ infant receiver mutual gaze onsets. D) Result of cluster-based permutation procedure for infant occipital ITC relative to adult sender/infant receiver non-mutual gaze onsets. For each, the top left plot is the raw differences where hotter (more yellow) colours indicate more ITC for looks to mutual vs non-mutual. The top right plot shows z scored

differences. The bottom left plot shows thresholded differences (at p=0.05) before cluster correction for multiple comparisons. The bottom right shows the final results of cluster-based permutation statistics after correction for multiple comparisons.

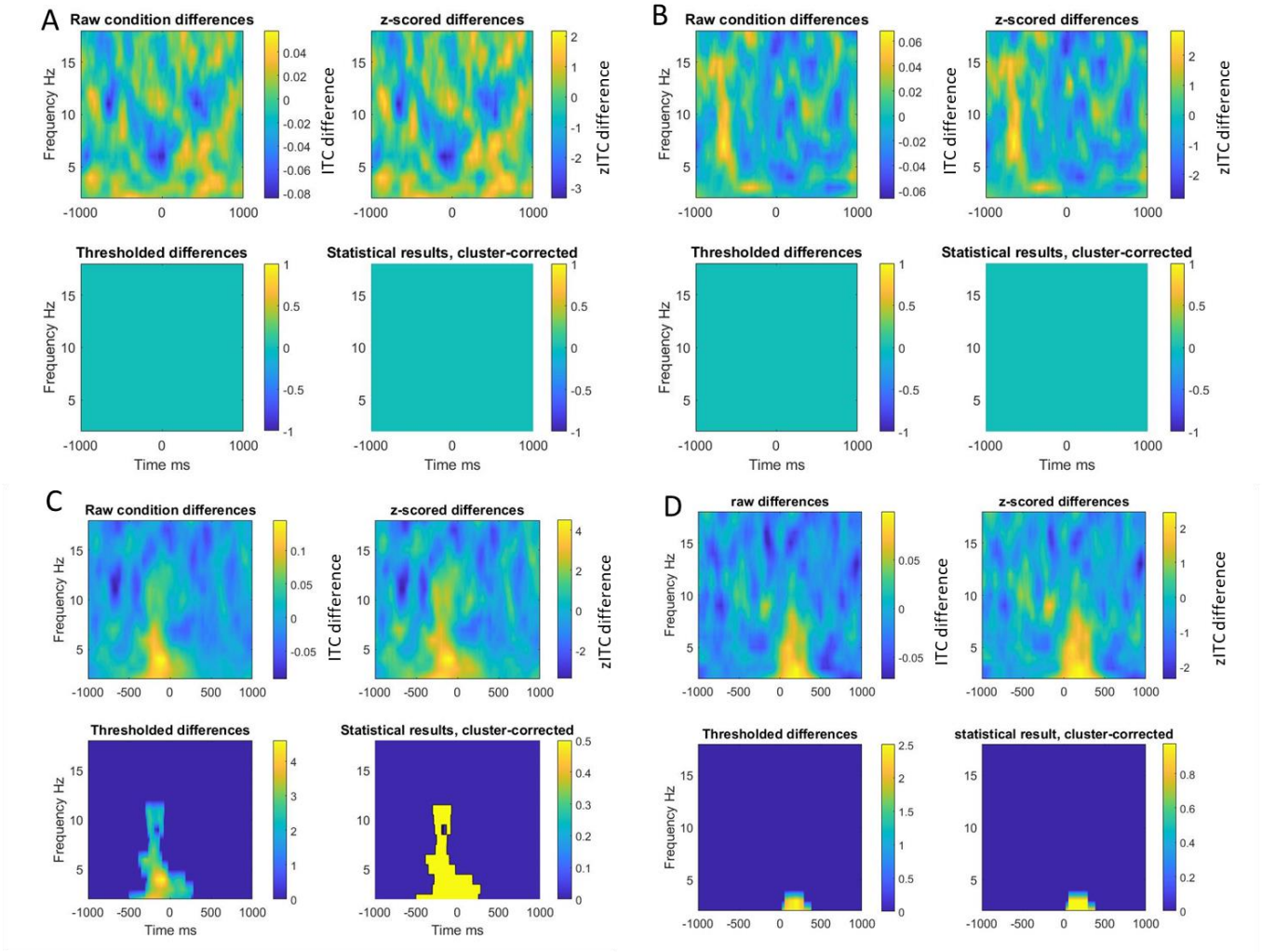


Figure 4.13. Result of cluster-based permutation procedure for event locked ITC relative to baseline. A) Result of cluster-based permutation procedure for adult occipital ITC relative to infant sender/ adult receiver mutual gaze onsets. B) Result of cluster-based permutation procedure for adult occipital ITC relative to infant sender/ adult receiver non-mutual gaze onsets. C) Result of cluster-based permutation procedure for adult occipital ITC relative to adult sender/ infant receiver mutual gaze onsets. D) Result of cluster-based permutation

procedure for adult occipital ITC relative to adult sender/ infant receiver non-mutual gaze onsets. For each, the top left plot is the raw differences where hotter (more yellow) colours indicate more ITC for looks to mutual vs non-mutual. The top right plot shows z scored differences. The bottom left plot shows thresholded differences (at p=0.05) before cluster correction for multiple comparisons. The bottom right shows the final results of cluster-based permutation statistics after correction for multiple comparisons.

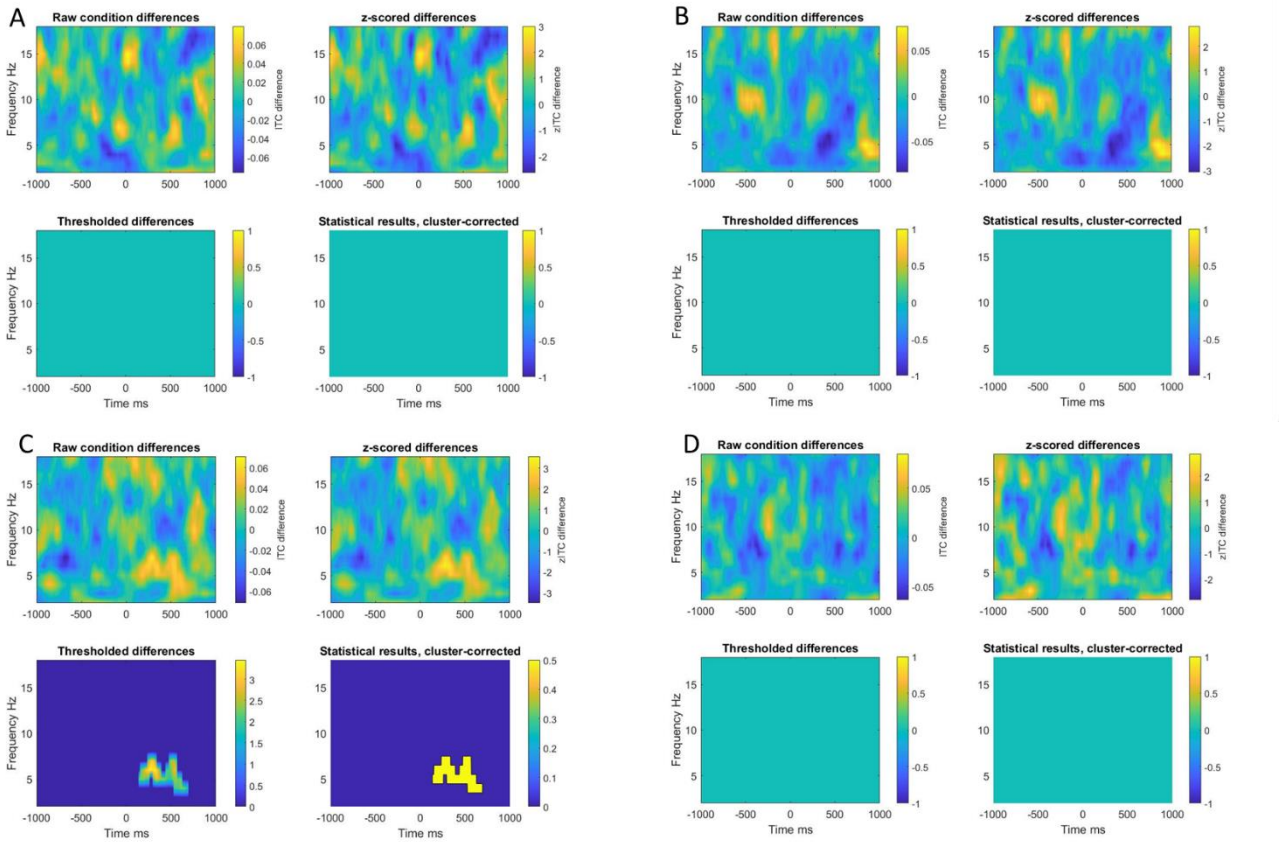


Figure 4.14. Result of cluster-based permutation procedure for event locked ITC for effect of gaze type A) Result of cluster-based permutation procedure for infant occipital ITC comparing infant sender mutual vs non-mutual gaze onsets. B) Result of cluster-based permutation procedure for adult occipital ITC comparing infant sender mutual vs non-mutual gaze onset. C) Result of cluster-based permutation procedure for infant occipital ITC

comparing adult sender mutual vs non-mutual gaze onsets. D) Result of cluster-based permutation procedure for adult occipital ITC comparing adult sender mutual vs non-mutual gaze onsets. For each, the top left plot is the raw differences where hotter (more yellow) colours indicate more ITC for looks to mutual vs non-mutual. The top right plot shows z scored differences. The bottom left plot shows thresholded differences (at p=0.05) before cluster correction for multiple comparisons. The bottom right shows the final results of cluster-based permutation statistics after correction for multiple comparisons.

4.7.5. *Inter-brain event locked analysis PLV and PDC – Results of permutation procedure*

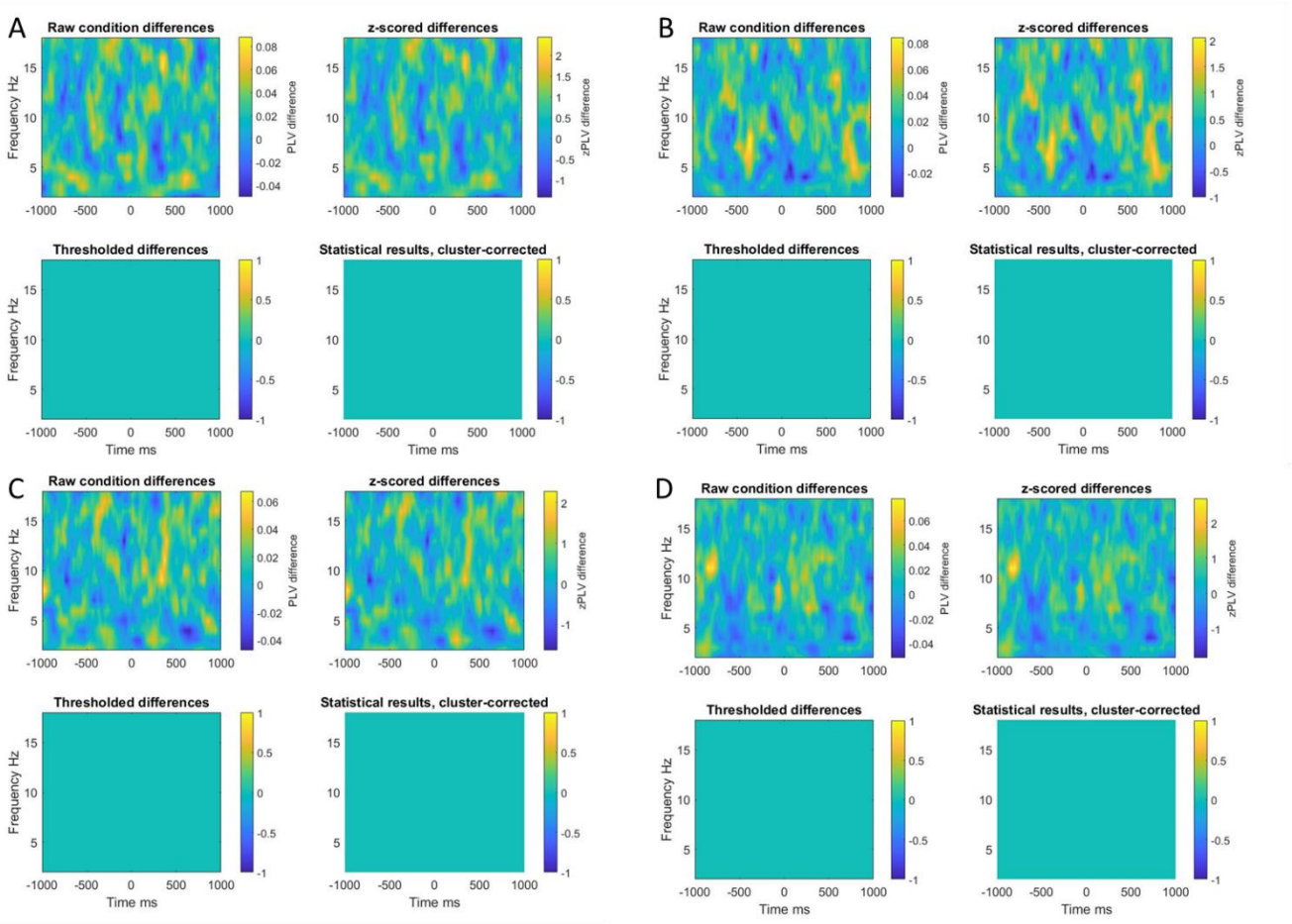


Figure 4.15. Result of cluster-based permutation procedure for event locked PLV relative to baseline. A) Result of cluster-based permutation procedure for occipital PLV relative to

infant sender mutual gaze onsets. B) Result of cluster-based permutation procedure for occipital PLV relative to infant sender non-mutual gaze onsets. C) Result of cluster-based permutation procedure for occipital PLV relative to adult sender mutual gaze onsets. D) Result of cluster-based permutation procedure for occipital PLV relative to adult sender non-mutual gaze onsets. For each, the top left plot is the raw differences where hotter (more yellow) colours indicate more PLV for looks to mutual vs non-mutual. The top right plot shows z scored differences. The bottom left plot shows thresholded differences (at p=0.05) before cluster correction for multiple comparisons. The bottom right shows the final results of cluster-based permutation statistics after correction for multiple comparisons.

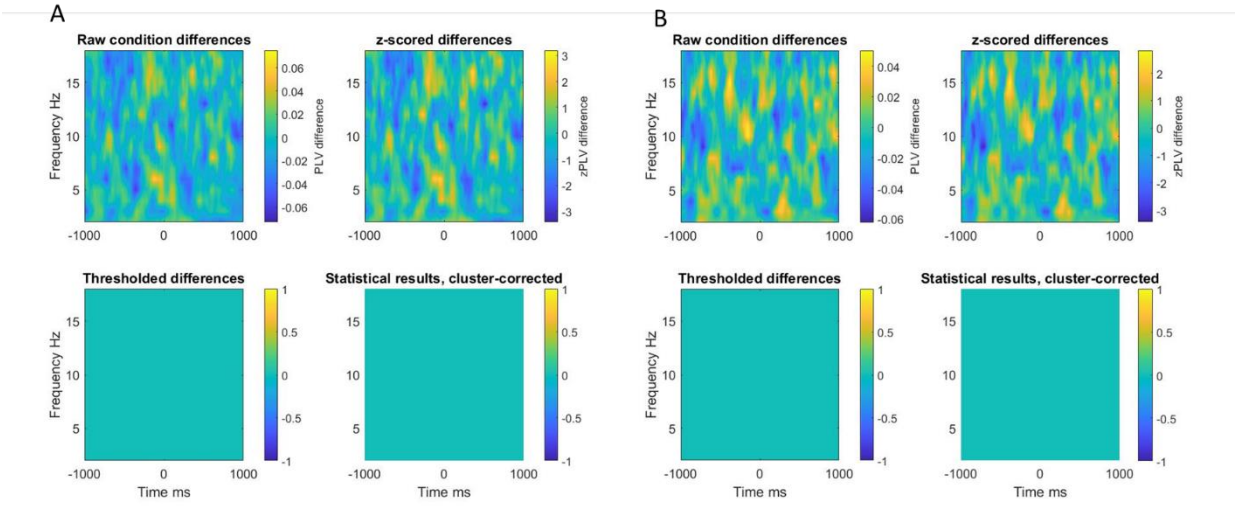


Figure 4.16. Result of cluster-based permutation procedure for PLV over occipital electrodes. A) shows difference between infant sender/ adult receiver looks to mutual vs non-mutual gaze onsets. B) shows difference between adult sender/ infant receiver mutual vs non-mutual gaze onsets (hotter, more yellow colours indicate more PLV for mutual vs non-mutual gaze).

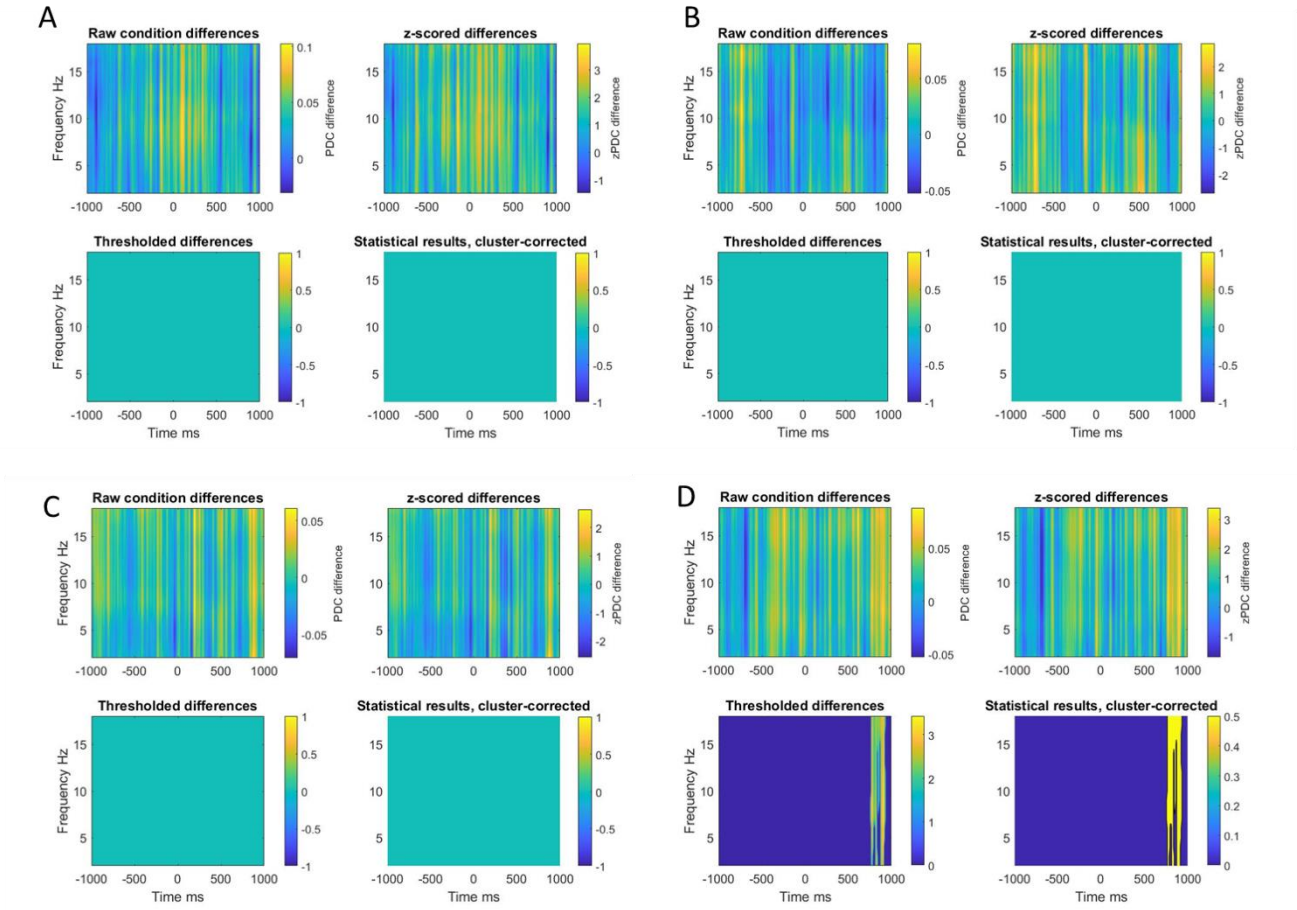


Figure 4.17. Result of cluster-based permutation procedure for event locked PDC relative to baseline. A) Result of cluster-based permutation procedure for occipital I→A PDC relative to infant sender/ adult receiver mutual gaze onsets. B) Result of cluster-based permutation procedure for I→A occipital PLV relative to infant sender/ adult receiver non-mutual gaze onsets. C) Result of cluster-based permutation procedure for occipital A→I PDC relative to infant sender/ adult receiver mutual gaze onsets. D) Result of cluster-based permutation procedure for A→I occipital PLV relative to infant sender/ adult receiver non-mutual gaze onsets. For each, the top left plot is the raw differences where hotter (more yellow) colours indicate more PDC for mutual vs non-mutual. The top right plot shows z scored differences. The bottom left plot shows thresholded differences (at p=0.05) before cluster correction for

multiple comparisons. The bottom right shows the final results of cluster-based permutation statistics after correction for multiple comparisons. I→A means infant to adult influence. A→I means adult to infant influence.

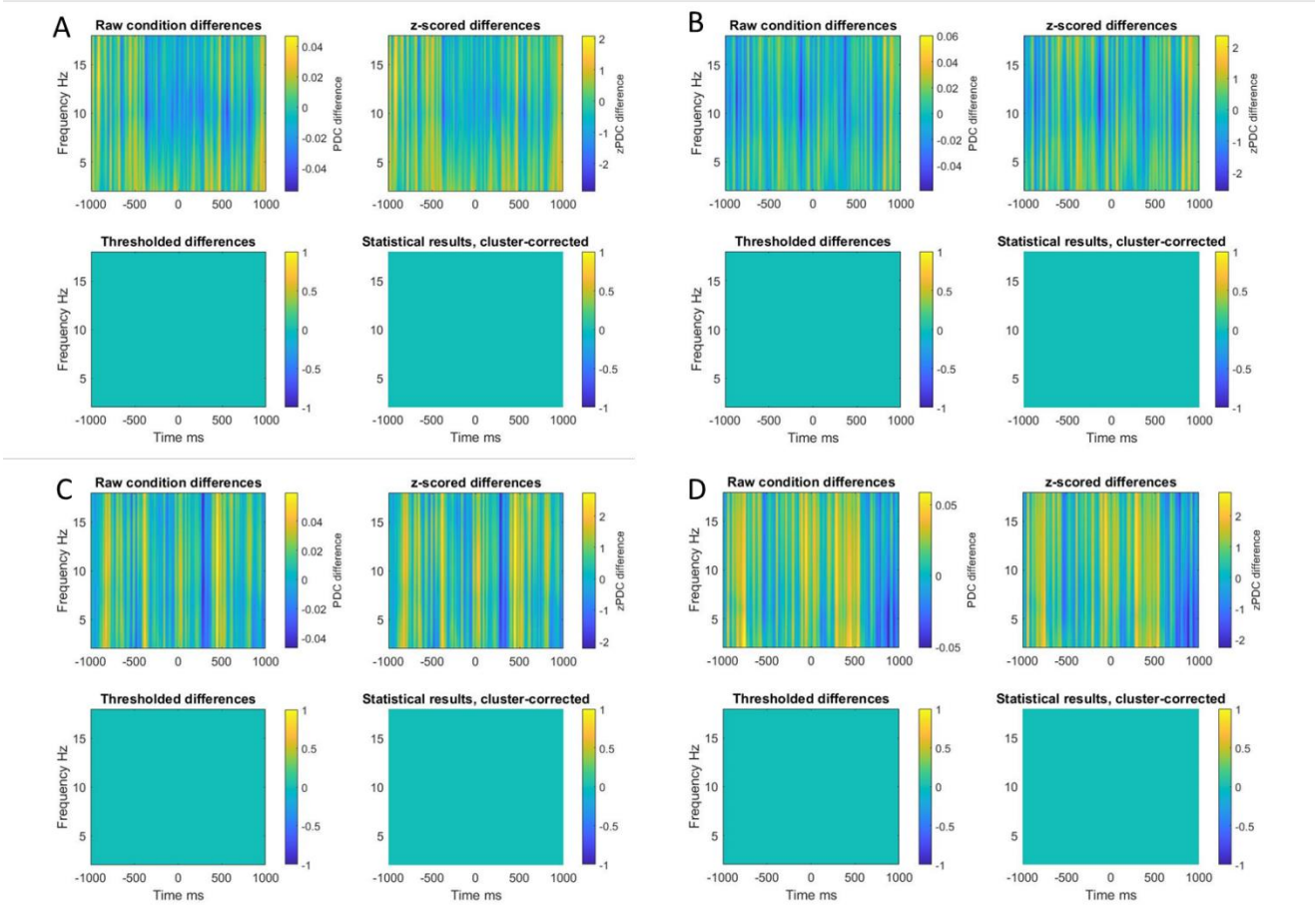


Figure 4.18. Result of cluster-based permutation procedure for event locked PDC relative to baseline. A) Result of cluster-based permutation procedure for occipital I→A PDC relative to adult sender/ infant receiver mutual gaze onsets. B) Result of cluster-based permutation procedure for I→A occipital PLV relative to adult sender/ infant receiver non-mutual gaze onsets. C) Result of cluster-based permutation procedure for occipital A→I PDC relative to adult sender/ infant receiver mutual gaze onsets. D) Result of cluster-based permutation procedure for A→I occipital PLV relative to adult sender/ infant receiver non-mutual gaze

onsets. For each, the top left plot is the raw differences where hotter (more yellow) colours indicate more PDC for mutual vs non-mutual. The top right plot shows z scored differences. The bottom left plot shows thresholded differences (at p=0.05) before cluster correction for multiple comparisons. The bottom right shows the final results of cluster-based permutation statistics after correction for multiple comparisons. I→A means infant to adult influence. A→I means adult to infant influence.

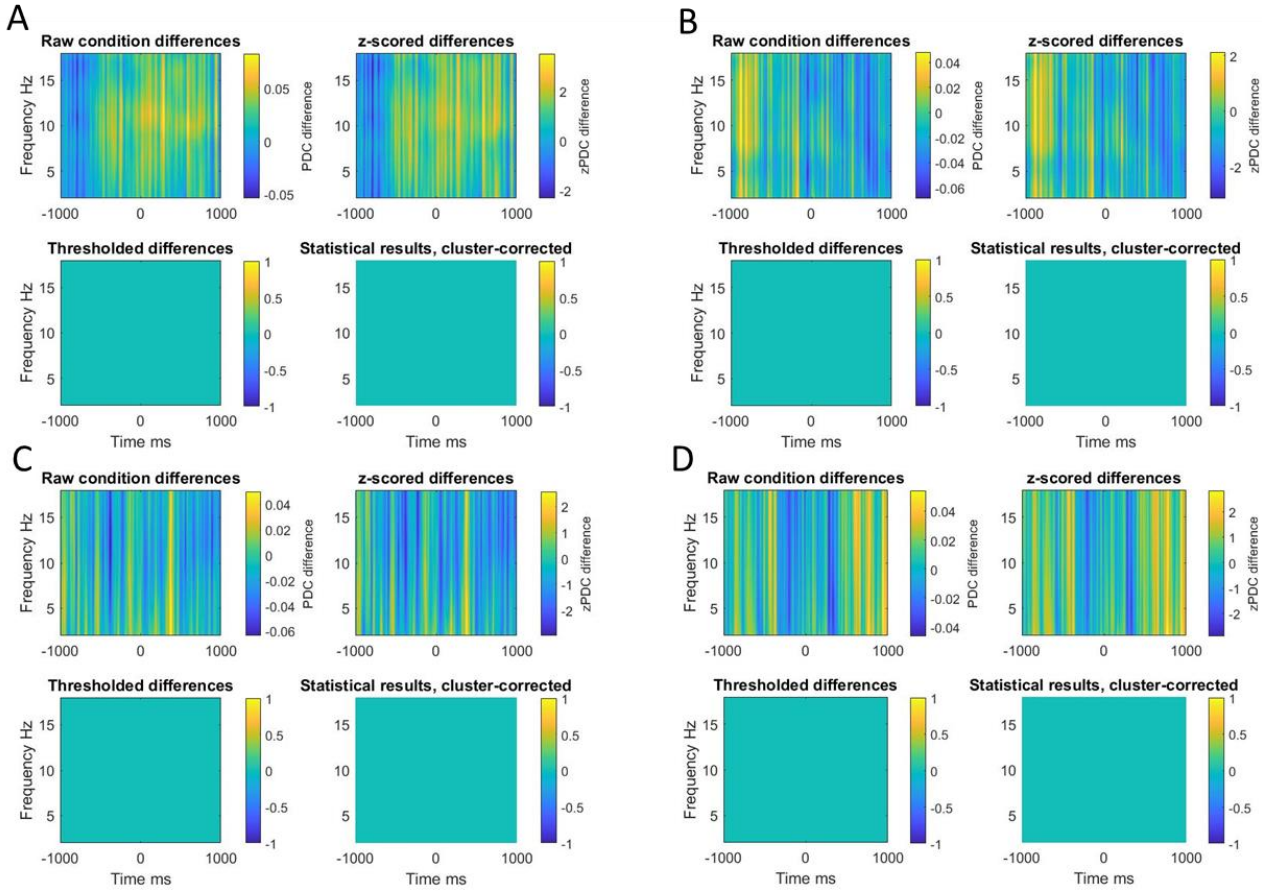


Figure 4.19. Result of cluster-based permutation procedure for PDC over occipital electrodes. A) shows difference between infant sender/ adult receiver mutual vs non-mutual gaze onsets for I→A influences. B) shows difference between infant sender/ adult receiver mutual vs non-mutual for A→I influence. C) shows difference between adult sender/ infant receiver mutual vs non-mutual gaze onsets for I→A influences. D) shows difference between

adult sender/ infant receiver mutual vs non-mutual gaze onsets for A→I influence (hotter, more yellow colours indicate more PLV for mutual vs non-mutual gaze). I→A means infant to adult influence. A→I means adult to infant influence.

4.7.6. Results of Bayes Factor analysis

4.7.6.1 Inter-brain non-event locked analysis – PLV and PDC

	Theta	Alpha
PLV	0.19/ 0.49/ 5.31	0.16/ 0.74/ 6.35
I→A PDC	0.16/ 0.70/ 6.22	0.16/ 0.66/ 6.08
A→I PDC	0.18/ 0.50/ 5.37	0.19/ 0.48/ 5.25

Table 4.1. Results of Bayes Factor analysis of inter-brain non-event locked PLV and PDC in theta and alpha. For each cell of the table the first number is the bf10 score, the second is the p value associated with the bf10 test and the third number is the bf01 score. These tests were conducted on average data over C3 and C4 electrodes in theta (3-6Hz) and alpha (6-9Hz).

4.7.6.2. Inter-brain event locked analysis – PLV and PDC

	Theta	Alpha
PLV	0.24/ 0.34/ 4.24	0.37/ 0.17/ 2.69

I→A PDC	0.18/ 0.55/ 5.58	0.18/ 0.52/ 5.42
A→I PDC	0.18/ 0.58/ 5.71	0.18/ 0.57/ 5.67

Table 4.2. Results of Bayes Factor analysis of inter-brain event locked PLV and PDC in theta and alpha. For each cell of the table the first number is the bf10 score, the second is the p value associated with the bf10 test and the third number is the bf01 score. These tests were conducted on average data over occipital electrodes in theta (3-6Hz) and alpha (6-9Hz).

4.7.7. Intra-brain analysis – ERP topoplots

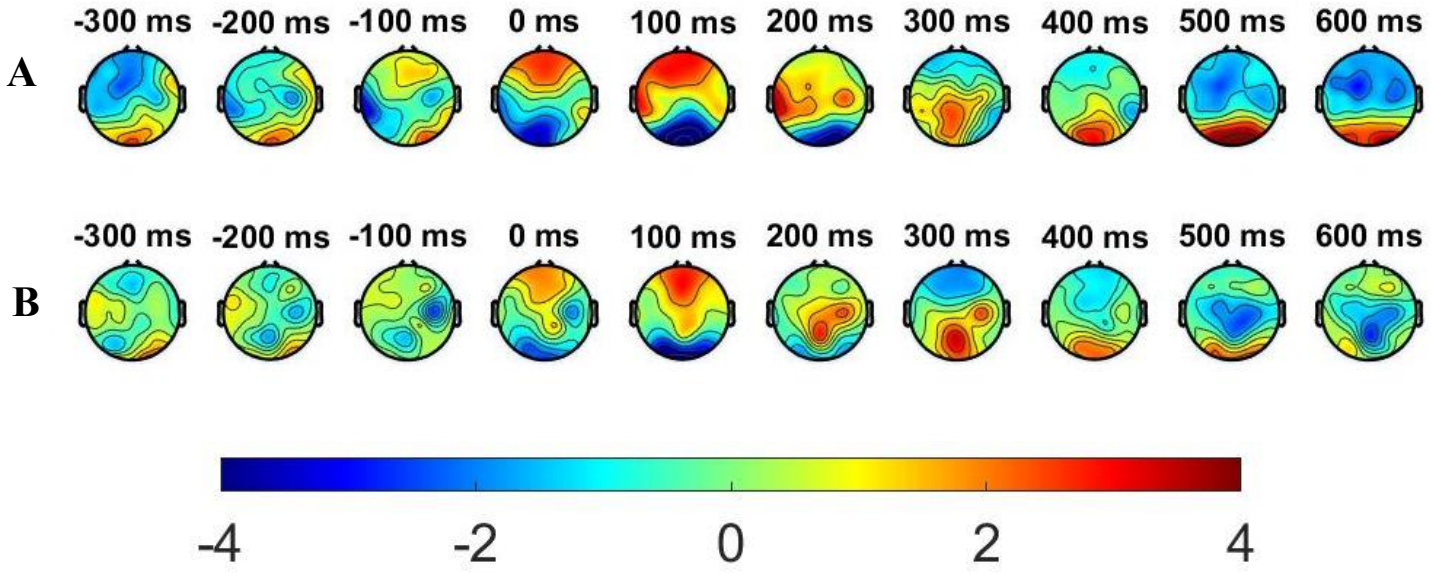


Figure 4.20. Topographical distribution of ERPs relative to infant sender mutual (A) and non-mutual (B) gaze onsets. Time 0 is onset of gaze.

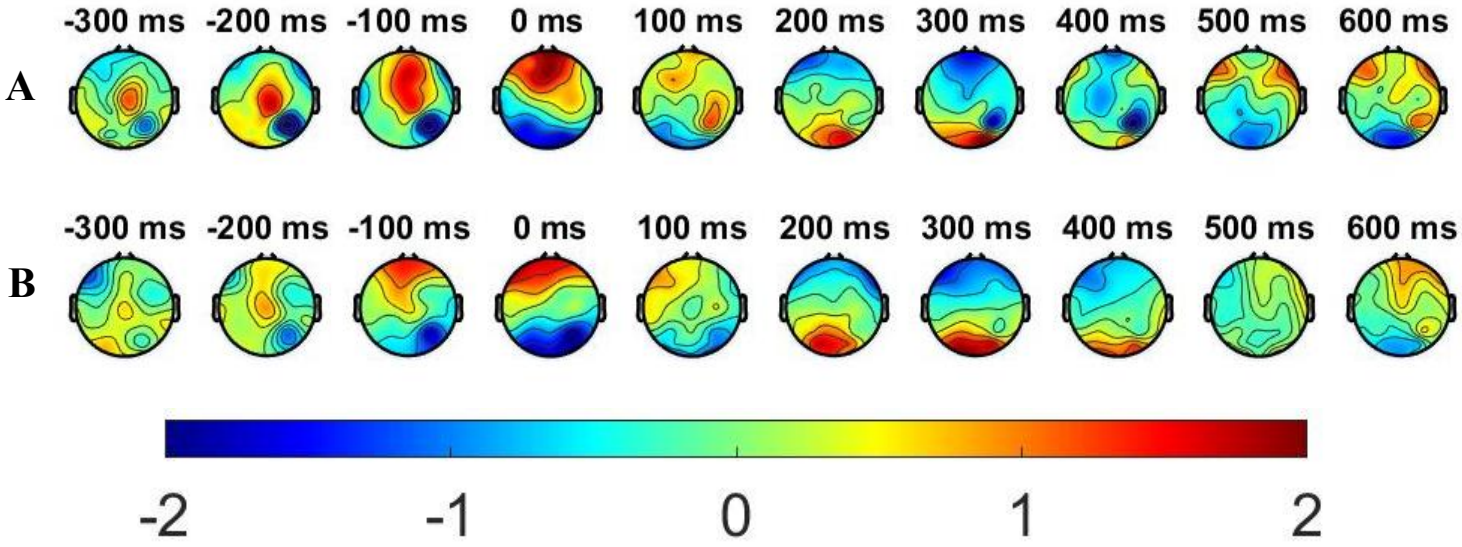


Figure 4.21. Topographical distribution of ERPs relative to adult sender mutual (A) and non-mutual (B) gaze onsets. Time 0 is onset of gaze.

4.7.8. Intra-brain analysis – ERP SEM

In this section we report the standard error to the mean of single trail and group average (GA) ERP data for infant and adult, sender and receiver, mutual and non-mutual gaze onsets.

	Sender						Receiver					
	Mutual			Non-mutual			Mutual			Non-mutual		
Sub no.	P1	N290	P400	P1	N290	P400	P1	N290	P400	P1	N290	P400
1	1.98	2.04	1.79	2.59	6.15	3	2.15	1.84	2.11	1.31	1.23	1.19
2	1.81	3.39	1.54	2.26	2.17	1.78	2.15	2.21	2.07	1.94	2.48	1.72
3	5.05	10.82	11.54	5.64	5.82	4.56	4.93	8.43	4.99	4.05	5.04	8.1
4	2.2	2.81	1.44	2.21	3.44	2.68	1.18	1.25	1.04	1.66	2.03	1.75
5	4.4	1.82	6.82	7.85	7.43	3.72	2.5	3.59	3.16	3.77	3.24	2.81
6	1.46	2.99	2.44	1.09	1.62	0.96	1.65	1.07	1.46	1.1	1.22	0.99
7	4.95	4.79	3.14	2.26	6.98	6.33	5.09	3.09	3.53	9.05	3.73	5.92
8	1.55	1.66	1.58	2.03	2.44	0.65	4.6	2.24	2.6	1.31	1.33	1.05
9	2.57	2.38	2.54	3.76	3.12	2.82	4.58	4.9	6.2	2.45	2.75	2.63
10	1.82	1.74	1.65	2.62	3.93	3.73	3.74	2.13	2.99	2.6	1.59	1.96
11	7.63	5.73	3.72	6.62	7.9	7.56	8.42	3.19	2.5	3.68	3.31	1.95
12	2.35	2.94	2.76	4	3.97	3.82	5.26	4.46	2.65	2.37	1.62	2.17
13	1.01	1.14	0.95	0.56	1.53	4.88	1.56	1.42	1.75	0.83	0.93	0.74
14	1.11	1.07	1.28	3.78	4.89	7.63	5.32	4.83	2.35	1.19	1.5	1.47
15	2.87	3.24	1.8	3.22	3.36	2.75	2.9	2.82	1.78	2.14	3.62	1.82
16	8.83	6.79	8.61	3.04	5.1	4.27	2.24	5.12	2.92	3.33	5.45	4.37
17	2.5	2.31	2.1	1.45	3.15	1.92	3.3	3.26	2.51	1.48	1.35	1.21
18	2.29	2.59	2.16	3.01	2.47	3.67	2.2	1.97	2.2	2.16	2.06	1.73

19	2.3	2.77	2.36	2.34	3.51	1.7	2.39	5.35	3.53	3.58	4	2.91
20	3.11	3.83	3.4	2.14	3.45	2.28	1.59	1.9	1.67	2.04	2.11	1.66
21	2.71	2.09	1.89	3.87	4.14	3.58	3.05	2.84	2.55	2.19	1.7	1.83
22	1.83	1.6	1.45	2.41	4.4	3.45	1.4	1.37	1.37	1.11	0.82	1.04
23	14.35	16.39	19.69	6.31	10.15	6.45	9.49	16.96	7.15	15.05	11.42	15.51
24	3.6	2.94	2.44	7.09	3.22	1.98	5.55	4.46	4.06	2.93	2.58	3.25
25	2.5	2.25	2.56	3.8	3.2	3.13	3.27	3.78	3.62	1.65	1.61	1.52
26	2.44	2.32	1.98	3.21	4.3	4.52	2.72	3.66	3.18	2.4	2.74	2.54
27	3.67	3.79	2.83	6.46	4.33	4.24	1.91	2.98	3.25	2.35	2.68	1.94
28	3.28	3.24	3.16	4.13	4.17	2.86	4.14	4.96	2.65	4.57	3.01	2.76
29	6.52	18.05	13.79	37.63	39.29	19.09	21.88	53.31	3.76	8.6	8.95	12.09
30	2.37	3.05	2.44	4.28	4.37	2.89	5.78	10.94	7.02	2.45	3.07	2.02
31	2.23	7.21	4.89	6.51	6.19	3.3	3.5	4	4.62	1.81	1.97	2.02
32	11.97	12.97	5.53	5.85	8.49	4.76	4.28	5.87	4.46	7.04	7.38	3.38
33	1.84	2.05	1.97	5.54	5.53	3.14	3.35	2.93	3.3	1.94	2.43	2.4
34	1.92	2.36	1.97	2.81	3.38	1.88	2.74	1.47	1.86	1.15	1.04	0.97
35	4.86	6.43	4.08	3.31	3.73	4.56	4.33	4.98	4.38	5.29	5.87	4.24
36	3.1	3.05	2	3	3.77	2.88	3.54	4.41	3.28	3.8	3.03	2.62
37	2.35	3.2	4.43	4.71	1.75	2.18	3.87	8.69	3.87	3.42	2.19	2.03
38	9.08	2.2	2.43	5.92	3	3.95	6.16	3.76	3.77	1.85	2.5	2.19

39	1.99	2.42	1.82	2.17	4.42	6.13	2.62	2.23	4.29	2.19	2.22	2.05
40	2.62	4.82	2.25	2.69	10.19	4.85	2.45	3.75	2.98	5	4.61	3.73
41	8.1	9.09	5.57	14.4	9.59	13.46	11.24	15.98	13.22	2.28	2.4	2.67
42	3.06	4.02	3.87	3.81	3.68	4.82	2.99	2.96	3.34	1.8	1.4	1.32
43	2.3	2.17	2.51	2.31	1.87	2.66	2.18	1.51	1.8	1.18	1.25	1.06
44	3	2.42	2.98	5.13	3.67	1.95	2.63	3.79	2.61	1.41	1.39	1.61
45	10.76	10.8	13.12	4.05	5.37	5.21	7.18	5.6	7.59	4.71	4.46	4.41
46	10.23	20.42	24.94	120.06	99.47	21.62	42.25	47.59	44.93	24.03	19.19	16.44
47	2.26	4.19	2.47	3.46	7.14	4.18	2.96	2.02	2.34	1.28	1.95	1.13
48	3.86	3.43	2.58	3.81	3.85	1.52	3.09	4.3	1.64	4.42	8.59	4.53
49	13.84	15.65	7.94	4.33	6.03	5.47	5.26	4.96	5.87	13.61	21.64	12.72
50	1.39	2.28	1.67	2.58	1.82	1.98	2.48	2.52	2	2.64	4.15	2.35
51	2.22	3.39	1.99	4.15	3.1	3.2	2.56	1.61	1.85	2.45	1.85	1.8
52	36.87	34.8	37.67	8.41	6.63	7.99	15.59	17.34	19.96	10.8	21.13	8.14
53	12.28	8.99	13.67	10.76	2.85	13.6	14.03	17.64	30.24	8	5.63	5.71
GA	1.26	2.23	3.32	5.22	2.12	2.41	1.26	0.9	3.32	0.99	2.12	1

Table 4.3. Results of standard error of the mean tests of infant ERP data. Note that removing outliers did not impact the significance of any of the tests, i.e., no tests that were not significant before became significant after removing outlier data a re-testing.

	Sender						Receiver					
	Mutual			Non-mutual			Mutual			Non-mutual		
Sub no.	P1	N170	P300	P1	N170	P300	P1	N170	P300	P1	N170	P300
1	284.78	291.87	196.88	359.57	227.26	338.18	307.82	198.98	269.4	273.98	108.67	175.12
2	111.08	121.85	109.46	133.38	105.24	196.23	106.21	162.83	115.62	206.06	177.56	164.46
3	1.25	1.29	1.24	2.14	1.78	2.28	1.86	1.62	1.37	1.22	1.53	1.12
4	0.62	0.7	0.69	0.57	0.9	1.22	0.41	0.45	0.49	0.49	0.64	0.61
5	1.83	2.81	1.89	3.22	4.14	2.71	3.42	2.96	2.53	1.93	2.93	2.51
6	0.48	0.64	0.42	0.49	0.49	0.42	0.7	0.76	0.41	0.56	0.74	0.46
7	1.22	0.87	1.05	1.71	2.48	1.39	1.94	2.46	1.59	1.83	2.19	2.16
8	4.27	3.05	3.83	6.89	10.03	9.64	13.5	6.2	4.52	1.74	1.86	2.29
9	1.1	1.23	1.1	1.35	1.21	1.15	0.91	1.08	0.99	1.68	1.59	1.25
10	2.34	1.92	2.97	4.45	4.05	4.82	3.85	4.67	3.27	2.27	2.28	2.74
11	4.52	1.72	3.52	3.29	3.82	3.28	6.72	6.03	5.02	5.02	1.23	2.88
12	1.01	0.96	0.94	2.12	3.59	1.74	1.21	1.67	2.55	1.15	1.06	0.96
13	1.94	1.09	1.6	4.08	3.94	5.91	1.69	1.28	1.91	0.77	0.76	0.8
14	0.93	1.23	0.82	4.05	3.9	1.15	2.67	3.42	2.54	0.89	1.04	0.93
15	1.1	1.24	0.94	2.95	2.62	1.19	1.17	1.19	1.44	0.96	1.41	0.95
16	1.23	1.07	0.9	1.86	1.34	1.22	2.6	2.25	1.71	1.12	1.46	1.33

17	1.16	1.25	1.24	1.06	1.09	1.28	1.61	1.81	1.38	1.48	1.38	0.97
18	5.92	3.34	5.41	1.36	1.65	2.11	8.54	4.28	4.35	0.77	1.1	0.77
19	0.66	0.5	0.62	6.28	7.91	23.3	2.59	1.84	1.43	7.73	5.22	4.38
20	1.45	2.04	1.44	0.79	0.67	1.44	1.24	1.27	0.92	0.72	1.21	0.8
21	2.35	3.29	2.29	1.53	1.72	2.43	1.69	1.52	1.82	1.14	1.13	0.92
22	2.5	2.86	2.09	1.32	1.31	1.54	1.75	3.24	1.68	1.11	0.71	0.81
23	1.32	2.34	1.56	4.52	4.46	3.75	3.95	6.49	1.78	2.58	2.05	2.03
24	3.66	2.87	2.58	2.45	5.31	5.23	5.72	3.79	3.72	2.24	4.37	2.93
25	1.27	1.26	1.3	0.62	0.94	0.23	1.84	2.05	1.81	6.37	3.23	2.34
26	4.38	2.95	2.5	2.62	1.54	2.14	2.37	2.57	2.02	0.73	0.82	0.73
27	1.28	1.35	1.52	2.92	3.95	5.87	1.78	1.63	1.97	3.29	2.31	2.85
28	1.21	1.13	1.04	2.43	2.54	2.95	2.18	2.14	3.01	0.89	1.03	0.77
29	2.88	2.67	1.94	1.1	1.86	1.26	5.16	0.35	4.02	1.54	1.2	1.68
30	1.56	3.74	4.32	9.02	12.08	6.15	4.03	2.08	2.79	4.53	2.02	2.72
31	4.57	1.82	4.4	1.13	1.15	0.96	2.75	1.58	3.69	0.97	1.06	0.94
32	3.16	3.27	2.6	2.28	0.96	2.14	5.56	3.65	3.98	1.65	1.99	1.87
33	0.9	1.21	0.86	2.82	2.77	2.72	1.67	1.31	1.58	2.05	1.35	1.37
34	1.1	1.14	0.91	3.42	4.03	1.94	1.15	1.05	1.04	1.23	1.45	0.91
35	1.49	1.31	1.57	1.44	1.4	1.53	1.48	1.91	2.04	0.62	0.61	0.71
36	0.46	0.62	0.68	2.16	1.94	2.3	0.78	1.55	0.49	1.16	1.07	0.75
37	1.3	1.52	1.01	0.6	0.38	0.46	1.46	1.15	1.19	0.4	0.37	0.39
38	0.9	0.94	1.02	0.85	0.93	0.93	1.64	1.8	1.05	1.04	1.42	1.59
39	1.69	1.15	1.5	2.84	1.56	1.41	1.58	2.28	1.56	1.36	1.24	1.09

40	4.23	2.62	3.22	1.75	1.95	1.46	5.92	4.09	4.2	0.81	0.8	0.83
41	1.57	1.59	1.37	29.09	13.31	10.47	4.59	2.41	1.77	9.36	5.36	9.08
42	7.08	10.04	8.6	4.14	4.16	1.94	11.06	12.77	11.81	1.35	1.2	0.94
43	1.09	1.42	1.07	28.07	17.53	29.14	1.3	1.38	1.35	6.61	6.48	6.07
44	1.72	1.2	1.16	1.57	0.92	1.94	1.48	1	0.83	1.14	0.94	0.98
45	1.66	3.78	2	1.73	1.61	1.43	7.08	5.67	6.15	0.84	0.77	0.54
46	3.18	4.36	6.74	8.67	5.45	2.12	4.98	3.02	3.41	1.93	1.82	1.36
47	41.98	56.28	46.06	2.91	3.32	0.76	64.97	73.52	90.82	2.05	1.46	2.75
48	8.23	2.69	8.32	95.69	48.11	72.56	3.73	3.45	2.96	35.5	44.77	55.47
49	2.55	2.6	1.86	9.46	5.74	7.92	1.77	2.02	2.3	2.69	2.33	3.06
50	1.9	1.13	1.61	3.12	3.58	4.35	2.42	1.31	1	2.07	10.36	1.66
51	209.04	131.15	200.34	6.81	2.33	5.28	569.93	958.7	624.33	2.79	5.34	4.01
52	57.3	23.1	42.75	1887.85	486.15	2090.58	24.53	10.17	15.13	446.77	1148.22	1181.83
53	1.52	0.74	1.41	16.55	7.24	17.88	0.35	0.92	4.59	3.74	5.67	1.3

G A	11.13	68.16	7.7	28.94	6.17	41.34	15.79	17.11	11.63	19.6	15.8	37.39

Table 4.4. Results of standard error of the mean tests of adult ERP data. Note that removing outliers did not impact the significance of any of the tests, i.e., no tests that were not significant before became significant after removing outlier data and retesting

4.7.9. Analysis of infant-adult dyads

To address the possibility that the discrepancy between the findings of the current study and those in the study by Leong and colleagues (2017) could be due to differences in partner familiarity, we performed an additional analysis with data collected from dyadic interactions in which the same infants interacted with a previously unfamiliar adult (a research assistant). This dataset included 17 infant-adult dyads (average 12m infants). This data was taken from visit 2 of a broader research program in which infants interacted with a research assistant (one of two). In visit 1 data for the main findings was collected, whilst infants interacted with their caregivers. The exact same experimental set up and paradigm was used for both visits, details of which can be found in section 4.2.4 of the main text. The results of this additional analysis are focused on our core research questions a) do we observe above-chance inter-brain synchrony during mutual gaze/ around mutual gaze onsets during free-flowing natural social interactions. b) do we observe phase resetting to mutual gaze onsets in natural contexts. c) if we observe above chance inter-brain synchrony and phase resetting around mutual gaze onsets, are they inter-related? Here, we wanted to investigate whether any of these were influenced by partner familiarity.

4.7.10. Intra-brain event locked analysis – ITC

Consistent with our main findings with infant-caregiver dyads (see section 4.4.4. of main text) we found that for infant-adult ITC values, over occipital electrodes, and frequencies 2-18 Hz, did not significantly exceeded baseline values generated from a permutation test. The permutation analysis indicated that ITC in the post-gaze onset window was significantly higher than baseline for sender mutual (Figure 4.22 A and D) and non-mutual (Figure 4.22 E and H) gaze onsets, in both parents and infants, but not for receiver mutual (Figure 4.22 B and C) and non-mutual (Figure 4.22 F and G) gaze onsets in either parents or infants. We then tested whether ITC values for the infant-adult dyads were greater for mutual vs non-mutual gaze onsets. The results of the cluster-based permutation analysis indicated no statistically significant differences between looks to mutual vs non-mutual gaze in the sender's brain activity in either adults or infants.

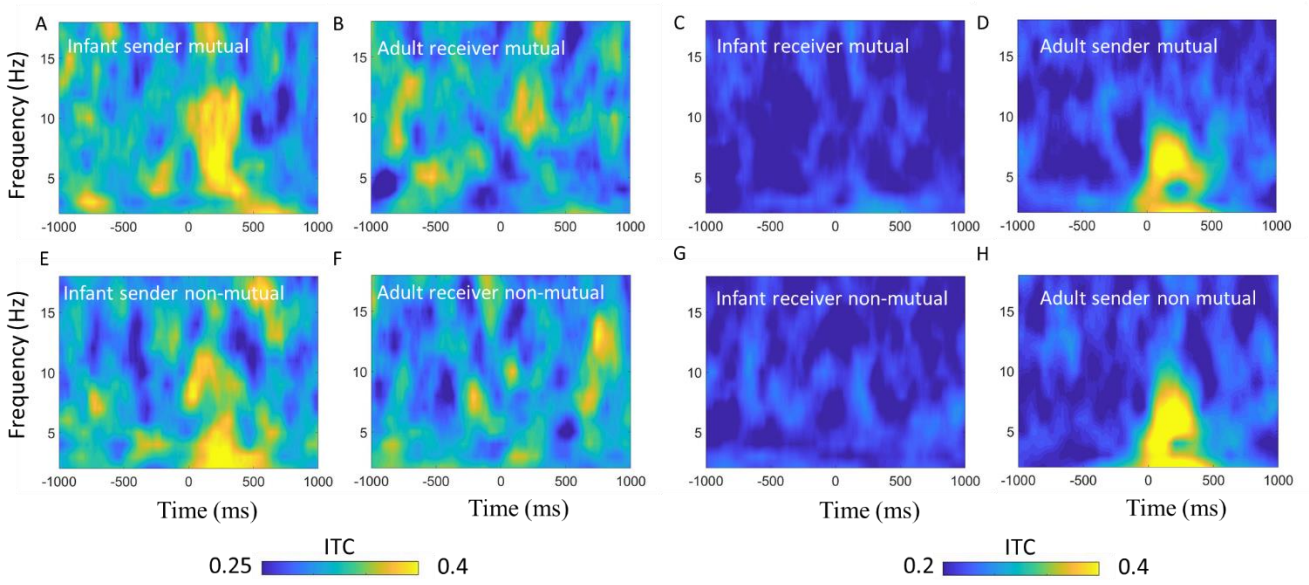
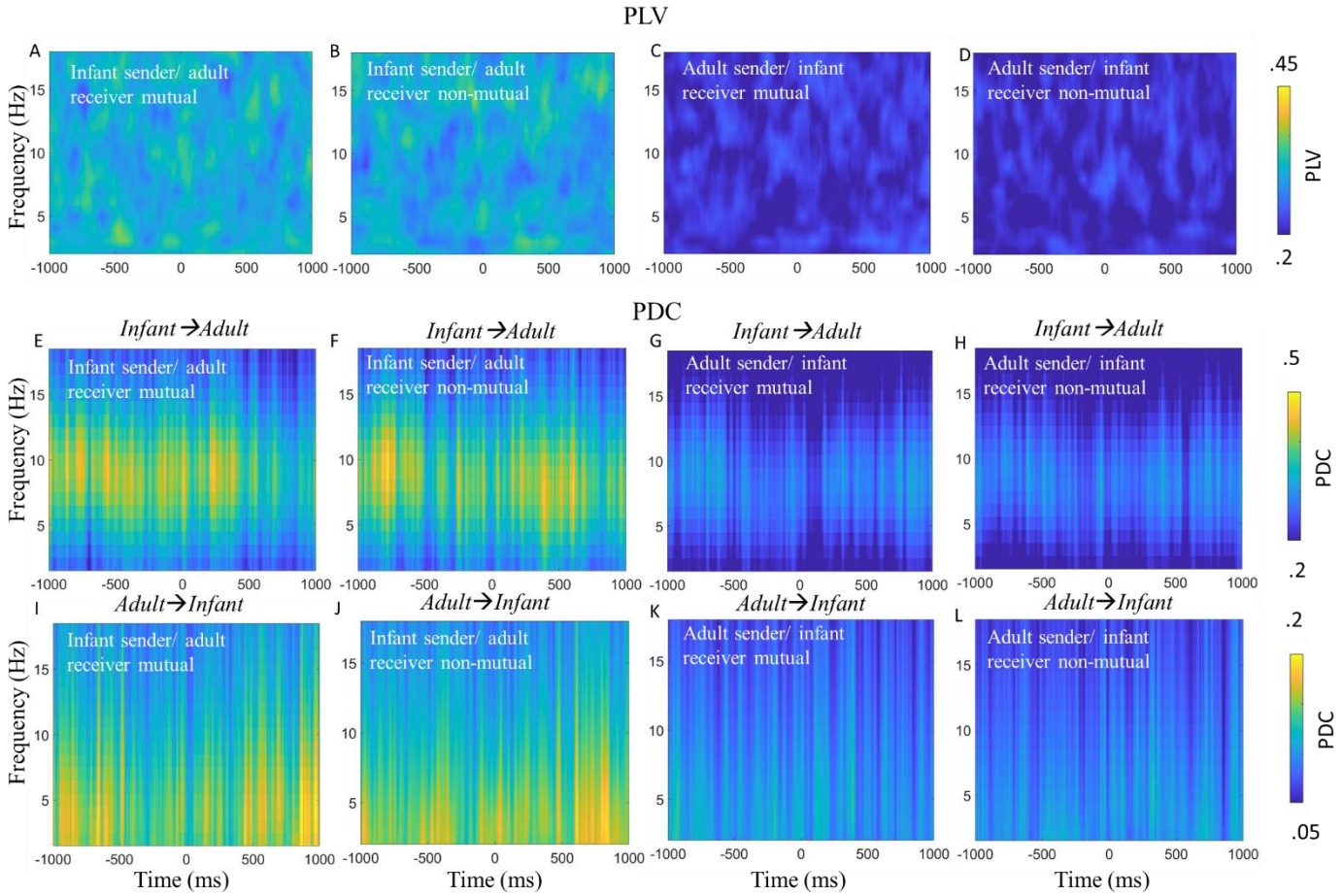


Figure 4.22. Inter-trial phase coherence time-locked to naturally occurring mutual and non-mutual gaze onsets. A) Infant ITC relative to infant sender mutual gaze onsets. B) Adult ITC

relative to adult receiver mutual gaze onsets. C) Infant ITC relative to infant receiver mutual gaze onsets. D) Adult ITC relative to adult sender mutual gaze onsets. E) Infant ITC relative to infant sender non-mutual gaze onsets. F) Adult ITC relative to adult receiver non-mutual gaze onsets. G) Infant ITC relative to infant receiver non-mutual gaze onsets. H) Adult ITC relative to adult sender non-mutual gaze onsets. No significant differences were identified.

7.7.11. Inter-brain event locked analysis – PLV

Again, consistent with our main findings with infant-caregiver dyads we found that PLV and PDC values over occipital electrodes and in the 2-18 Hz range did not significantly exceed baseline values generated from a permutation procedure for infant sender/ adult receiver and



adult sender/ infant receiver looks to mutual and non-mutual gaze. The result of the permutation analysis indicated that event-locked PLV and PDC values around mutual and non-mutual gaze onsets were not significantly different from baseline values. We also observed no statistically significant differences for the effect of gaze type (e.g., mutual vs non-mutual).

Figure 4.23. Infant-caregiver inter-brain synchrony time-locked to naturally occurring mutual and non-mutual gaze onsets. A) PLV relative to infant sender/adult receiver mutual gaze onsets. B) PLV relative to onsets of infant sender/adult receiver looks to non-mutual gaze. C) PLV relative to adult sender/infant receiver mutual gaze onsets. D) PLV relative to adult sender/infant receiver non-mutual gaze onsets. E) Infant→Adult PDC relative to infant sender/adult receiver mutual gaze onsets. F) Infant→Adult PDC relative to infant sender/adult receiver non-mutual gaze onsets. G) Infant→Adult PDC relative to adult sender/infant receiver mutual gaze onsets. H) Infant→Adult PDC relative to adult sender/infant receiver non-mutual gaze onsets. I) Adult→Infant PDC relative to infant sender/adult receiver mutual gaze onsets. J) Adult→Infant PDC relative to infant sender/adult receiver non-mutual gaze onsets. K) Adult→Infant PDC relative to adult sender/infant receiver mutual gaze onsets. L) Adult→Infant PDC relative to adult sender/infant receiver non-mutual gaze onsets. No significant differences were identified.

4.7.12. Analysis of non-event locked inter-brain synchrony using different pre-processing procedures

To address the possibility that the discrepancy between the findings of the current study and those in the study by Leong and colleagues (2017) could be due to differences in pre-processing procedures, we cleaned our data following to the best of our ability the pre-processing procedures outlines in the study by Leong and colleagues (2017). These steps were as follows. First the data were filtered between 0.1 and 100Hz and references to the vertex (Cz) electrode. Second sections where the baby was inattentive were removed, marked by visual inattention to either their partner or object. Third datapoints with an amplitude above 100mv were removed. Fourth the data was down sampled from 512Hz to 256Hz. Finally, the data was low passed filtered at 45Hz. Following this we then ran the exact same analysis procedures detailed in section 4.3.1 of the main text. We then conducted a repeated measures (RM) ANOVAs using average indices (over C3 and C4), taking frequency (2 levels) and gaze type (mutual vs non-mutual; 2 levels) as within-subjects factors. The results of the ANOVA indicated no statistically significant differences in PDC; for infant to adult (*Infant*→*Adult*) influences, $F(1, 55) = 1.08$, $p = .3$ or adult to infant (*Adult*→*Infant*) influences, $F(1, 55) = 0.003$ $p = .96$, between mutual and non-mutual gaze. The results did indicate a significant effect of frequency (i.e., more synchrony in Theta than Alpha) for PLV, $F(1, 55) = 25$, $p < .01$ and PDC; *Infant*→*Adult*, $F(1, 55) = 41.2$, $p < .01$ and *Adult*→*Infant*, $F(1, 55) = 191.17$, $p < .01$).

Overall, the results of these analyses rule out the possibilities that the discrepancies between the findings of the current study and those in the study by Leong and colleagues (2017) are due to differences in partner familiarity or differences in pre-processing procedures.

Chapter 5 - Does multimodal behavioural synchrony associate with parent-infant inter-brain synchrony?

The following chapter summarises a study investigation the relationship between inter-brain and behavioural synchrony. Subheadings, figure placement, figure and table numbers, and citation style have been adapted to conform to the general thesis format.

Abstract

Social interactions involve complex sequences of prediction and response between interacting individuals that unfold simultaneously across a multitude of behavioural modalities and physiological levels. Despite a growing prevalence of theories suggesting increased behavioral synchrony during social interactions associates with elevated levels of inter-brain synchrony between interacting individuals, little is known about how this arises during development. Here we analysed patterns of inter-brain synchrony in parent-infant dyads during naturalistic social interactions (re-analysis of a N=55 parent-infant, approx. 12m old, dual EEG dataset). We took an exploratory data driven approach, investigating patterns of above chance inter-brain synchrony in topographical, frequency and temporal space using nonparametric statistics and cluster correction across the three dimensions. Once above chance patterns of inter-brain synchrony had been identified we were then able to investigate how different features (e.g., synchronised vs un-synchronised behaviour) and which modes (e.g., hands vs gaze) of behavioural coordination contributed to these patterns. We found consistent with previous research significant pattens of behavioural and inter-brain synchrony during naturalistic parent-infant social interactions. However contrary to our hypothesis we

found no associations between behavioural and inter-brain synchrony, but rather we observed the strongest associations were between EEG power and inter-brain synchrony, suggesting this was a greater driver of variation in inter-brain synchrony. Overall, our results further previous recent work suggesting that moments of shared visual attention do lead to increases in parent-infant inter-brain synchrony.

5.1. Introduction

Social interactions involve complex series of predictions and response between interacting individuals that unfold simultaneously across a multitude of behavioural modalities and physiological levels. Even from very early on in development infants and adults have an extensive repertoire of coordination sequences that are substantiated across a range of behavioural modalities (Yu and Smith, 2013). For example, recent studies (Franchak et al., 2011; Yu & Smith, 2013) have shown that infants rarely look to their caregiver's face and eyes during free-flowing interactions but achieve coordination with their caregiver through temporal and spatial tracking of hand movements.

In recent years there has been increasing focus on exploring how the brain activity of two individuals synchronises during social interactions. This research suggests that across a range of different social settings (e.g., ranging from inter-brain synchrony between multiple school aged children; Dikker et al., 2017 to infant-caregiver dyads interacting contingently through a screen; Reindl et al., 2022) and cognitive/ task demands the neural activity of two or more individuals becomes coupled. However, it remains an open question exactly how this is substantiated in the brain.

Whilst there may be additional higher order mechanisms, for example some research suggests that shared understanding is associated with shared neural responses between individuals (Simony et al., 2016), that contribute to inter-brain synchrony, most researchers agree that inter-brain synchrony during social interactions is embedded within interpersonal behavioural coordination (Dumas et al., 2011, Burgess 2013, Wass et al., 2020, Hamilton, 2021, Novembre & Iannetti, 2021). This is supported by a mounting body of research that has linked inter-brain synchrony with behavioural coordination across various modalities; a number of studies have observed increased inter-brain synchrony during moments of coordinated action (Dumas, 2010, Lindenberger et al., 2009, Yun et al., 2012, Sänger et al., 2013, Gamliel et al., 2021). For example, in their study of inter-brain synchrony during a task which required adult-adult dyads to synchronise their hand movements Gamliel and colleagues (2021) found, using fNIRS hyperscanning, both increased behavioural coordination and increased inter-brain synchrony for ingroup vs inter-group participants.

A number of studies have also found associations between inter-brain synchrony and mutual gaze. For example, in interacting adult-adult dyads Luft and colleagues (2022) found that mutual gaze was associated with higher inter-brain gamma band (30-45Hz) coherence (a spectral measure based on correlation) between interacting adults than non-mutual gaze. In the developmental literature, our group investigated inter-brain synchrony in 7.5-month infant-adult dyads during moments of mutual and non-mutual gaze (Leong et al., 2017). Taken together these studies imply that there is a positive relationship between behavioural coordination (although see Haresign et al., 2022b) during social interaction and inter-brain synchrony.

Although this evidence suggests that increased inter-brain synchrony is associated with moments of behavioural coordination very little is still known about the mechanisms that give rise to this (Liu et al., 2018; Haresign et al., 2022). One of the most commonly proposed mechanisms for establishing inter-brain synchrony is concomitant phase resetting of neural oscillations around moments of behavioural coordination (Wass et al, 2020; Leong et al., 2017; Luft et al., 2022). It is known that the phase of neuronal oscillations reflects the excitability of underlying neuronal populations to incoming sensory stimulation (Klimesch et al., 2007; Jensen et al., 2014) and that sensory information arriving during high-receptivity periods is more likely to be perceived than information arriving during low-receptivity periods (Busch et al., 2009; Mathewson et al., 2009; 2010; 2011; 2012), suggesting that there exist mechanisms (phase resetting) for modulating the phase of neuronal oscillations, in order to match the temporal structure of the input (van Diepen et al., 2015; Ruzzoli et al., 2019). Although this is one of the most popular ideas for how inter-brain synchrony arises during social interactions only one study has empirically tested this theory; In our recent work (see Haresign et al, 2022) we investigated whether onsets of mutual gaze led to concomitant phase resetting and increased inter-brain synchrony in (12m) infant-adult dyads engaged in naturalistic free play interactions. In that work we were unable to provide evidence linking phase resetting to moments of behavioural coordination and therefore we were unable to establish phase resetting as key mechanism for how inter-brain synchrony arises during social interaction. While for now the mechanisms supporting fine grained inter-brain synchrony remain unclear, even over larger timescales increased understanding of the basic characteristics of inter-brain synchrony would help to build more/better hypotheses for how it arises during social interaction. For example, how do patterns of interbrain synchrony propagate through topographical, frequency and time space? And which behavioural modalities are most salient?

Analysis of interbrain synchronisation using 3-dimension non-parametric statistics

To date most studies of inter-brain synchrony have chosen to analyse two-dimensional (2D) activity, using mass-univariate approaches, whereby inter-brain synchrony is computed for every possible paired data sample. Meaning that in these studies, often researcher driven decisions have been made as to which dimensions are restricted. For example, many studies analyse the whole topographical space, whilst restricting frequency to the canonical bands (e.g., Perez et al., 2017, Santamaria et al., 2020) or they analyse all time-frequency activity whilst restricting topographical information (Lindenberger et al., 2009, Haresign et al., 2022). This makes sense as the datasets analysed in these studies are typically quite large (n=55 infant-adult dyads analysed in Haresign et al., 2022) and inter-brain synchrony is computed over a sizeable number of data points. Therefore, even with 2D activity these analyses are computationally very intensive. However, making decisions about where to limit one's search can be problematic, particular within the field of inter-brain synchrony, given the overall lack of empirical research on which to base predictions about where and when an effect might be observed. Therefore, the field would benefit from increased focus on approaches that analyse the entire data space (i.e., in topographical, time and frequency space). These approaches have the merit of not having to choose locations, frequencies or time points a priori and therefore allows for potentially observing non-expected effects.

One way in which researchers could do this, would be through conducting nonparametric permutation tests, which estimate the distributions of a test static or data value that would likely be expected to occur within the data by chance. Overall nonparametric tests make no a priori assumptions about the data and therefore are well suited to exploratory analyses involving multiple dimensions and large numbers of data points. Here the distributions are

created from the data through procedures that iteratively shuffle some aspect of the data. For example, this could be the labels associated with two experimental conditions assigned to individual trails. Here the trials could be pooled together, and each trial randomly assigned one of two labels. The trials could then be split into the two conditions and compared by means of t-test. This is then repeated a number of times (often 1000) and a distribution of test statistics under the null hypothesis (in this example that there is no difference between conditions) is obtained. Next the same between condition comparison is performed with the real data with the original trial labels. Statistical evaluation works by comparing the observed test statistic for the real data with the distribution of test statistic created under the null hypothesis. Significance is declared if the observed t-value is an outlier, usually defined by the upper 5% of the distribution. This statistical procedure could then be repeated, and a test statistic/ *p*-value could be obtained for each channel time and frequency combination. The flexibility of nonparametric tests makes them ideal for exploratory analyses of inter-brain synchrony, which is why it is surprising that only very few studies within this field have utilised these methods (Dumas, 2010, Haresign et al., 2022, Gugnowska et al., 2022), a fact that has been the topic of recent criticism (Holroyd, 2022).

Overview of paper

Here, we analysed patterns of inter-brain synchrony in parent-infant dyads during naturalistic social interactions. We took an exploratory data driven approach, investigating patterns of above chance inter-brain synchrony in topographical, frequency and temporal space using nonparametric statistics and cluster correction across the three dimensions. This allowed us to first investigate how patterns of parent-infant inter-brain synchrony propagate through these various dimensions without constraining this analysis to specific behaviours. Once above

chance patterns of inter-brain synchrony were identified we were then able to investigate how different features (e.g., synchronised vs un-synchronised behaviour) and which modes of behavioural coordination contributed to these patterns. Our research questions were four-fold; Firstly, building on large body of research (e.g., Richardson et al., 2007) we looked at whether infant's and caregiver's gaze is coupled above chance during natural interactions. We predicted that parents and infants gaze and touch behaviours would be significantly cross correlated above chance. Second, building on previous work (Wass et al., 2018), we looked at whether infant's own looking behaviours are associated with intra-brain EEG power above chance during joint play with caregivers. Third, we looked at whether infant's intra-brain EEG power is correlated with joint attentional processes, through this we aimed to ask the question of what are the neural mechanisms in infancy that support joint attention and joint touch with a caregiver? Fourth we looked at whether joint patterns of gaze and touch associated with inter-brain synchrony.

5.2. Methods

5.2.1. Ethics statement

This study was approved by the Psychology Research Ethics Committee at the University of East London. Participants were given a £50 shopping voucher for taking part in the project.

5.2.2. Participants

Of the 90 infants we tested for this study, 21 contributed no data at all, 6 contributed EEG data that was too noisy even after data cleaning and 4 were lost due to human error, e.g.,

failed synchronisation triggers. We also excluded all participants with fewer than 5 gaze onsets, leading to an additional 4 datasets being excluded. The final sample contained 55 healthy (23 F), *M* = 12.2-month-old (*SD* =1.47) infants, that participated in the study along with their mothers.

5.2.3. Experimental set-up and procedure

Caregivers and infants were seated facing each other on opposite sides of a 65cm wide table. Infants were seated in a highchair, within easy reach of the toys. The shared toy play comprised two sections, with a different set of toys in each section, each lasting ~5 minutes 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, whilst leaving the table uncluttered enough for caregiver and infant gaze behaviour to be accurately recorded. 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. Both 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 minutes (SD=2.31).

5.2.4. Behavioural data

Video recordings were made using Canon LEGRIA HF R806 camcorders recording at 50fps positioned next to the infant and parent respectively. Video recordings of the play sessions were coded offline, frame by frame, at 50 fps. This equates to a maximum temporal accuracy of ~20ms.

5.2.5. 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. Coders indicated the start frame (i.e., to the closest 20ms, at 50fps) 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. Inter-rater reliability analysis on 10% of coded interactions (conducted on either play section 1 or play section 2), dividing data into 20ms bins, indicated strong reliability (McHugh, 2012) between coders (kappa=0.9 for caregiver coding and kappa=0.8 for infant coding).

The target of parent and infant visual attention and manual touch was coded separately by trained coders who manually coded frame-by-frame using custom built MATLAB scripts which synchronized the three video streams. A continuous, mutually exclusive, coding scheme was used, with the target of gaze coded as either toward *object 1*, *object 2*, *object 3*, *partner* (face only), *inattentive* or *uncodeable*; and the target of touch either *object 1*, *object 2*, *object 3*, *partner*, *nothing* or *uncodeable*. The target of touch was coded separately for each hand for both parent and infant. Frames were coded as *uncodeable* when parent/infant hands or eyes were obscured from view and when a researcher entered the experimental set-up.

Moments of joint touch were identified by synchronising parent infant touch data, which was then synchronised with the EEG data,

5.2.6. EEG data acquisition

EEG signals were obtained using a dual 32-channel Biosemi system (10-20 standard layout), recorded at 512 Hz with no online filtering using the Actiview software. EEG was time-locked to the behavioural data offline based on the video coding using synchronized LED and TTL pulses. To verify the synchronisation, we manually identified blinks in the behavioral data and looked to see if this matched the timing of the blinks in the EEG data.

5.2.7. EEG artifact rejection and pre-processing

A fully automatic artifact rejection procedure was adopted, following procedures from commonly used toolboxes for EEG pre-processing in adults (Mullen, 2012; Bigdely-Shamlo, et al., 2015) and infants (Gabard-Durham et al., 2018; Debnath et al., 2020). Full details of the pre-processing procedures can be found in (Haresign et al., 2021). In brief the data was filtered between 1 and 20Hz and re-referenced to a robust average reference. Then we interpolated noisy channels based on correlation; if a channel had a lower than 0.7 correlation to its robust estimate (average of other channels) then it was removed. The mean number of channels interpolated was 3.9 (*SD* =2.1) for infants and 3.9 (*SD* =4.4) for adults. Then for the infant data only we removed sections from the continuous data in which the majority of channels contained extremely high-power values. Data was rejected in a sliding 1 second epoch and based on the percentage of channels (set here at 70% of channels) that exceeded 5

standard deviations of the mean EEG power over all channels. For example, if more than 70% of channels in each 1-sec epoch exceeded 5 times the standard deviation of the mean power for all channels then this epoch is marked for rejection. We found that for adults this step was primarily removing activity that could be removed with ICA (e.g., eye movement artifact) without removing entire sections of the data. The average amount of continuous data removed was 11.9% (*SD* =14.6%) for infants. Finally, we used ICA to remove additional artifacts. Careful attention was paid to artifact and the amount of noise in the data throughout. In the supplementary materials we report the results of standard measures of EEG data quality (Luck et al., 2021). Including universal measures like these enables fast and easy comparison between studies and allows the overall quality of the data to be readily assessed.

5.2.8. Time frequency analysis- extracting power and phase

Time-frequency power and phase was extracted via complex Morlet wavelet convolution. The wavelets increased from 2 to 18 Hz in 17 linearly spaced steps and the number of cycles increased from 3-10 cycles logarithmically (this approach is generally recommended; see Cohen, 2014, chapter 13).

5.2.9. Inter-brain–PLV

For the non-event locked analyses, the phase locking value (PLV) was calculated within a single trial over a defined temporal window (e.g., Tass et al., 1998) according to:

$$PLV_t = \frac{1}{T}\left| \sum_{n=1}^{T} e^{i(\phi(t,n)-\psi(t,n))} \right| \quad (2),$$

Where T is the number of observations or time samples within the window, $\phi(t,n)$ is the phase on observation *n*, at time *t*, in channel ϕ and $\psi(t,n)$ at channel ψ. PLV across the entire interaction was computed in a 1-sec sliding window with 50% overlap between consecutive windows.

5.2.10. Inter-brain synchrony- PDC

Partial directed coherence (PDC) is based on the principles of Granger causality (Baccalá and Sameshima, 2001). It provides information of the extent to which one times series influences another. PDC is calculated from coefficients of autoregressive modelling according to:

$$PDC = \left(\frac{A_{xy}(f)}{\sqrt{a_y^*(f).a_x(f)}} \right) \quad (2),$$

where $A_{xy}(f)$ is the spectral representation of bivariate model coefficients and a_y and a_x are the spectral model coefficient from the univariate autoregressive model fit. Based on previous literature (e.g., Leong et al., 2017) we chose to compute PDC in 1-second non-overlapping sliding window. We estimated the model order for each segment using Bayesian information criteria (BIC). Model order values were then averaged for all segments. The result was a model order of 5, the same as used by Leong and colleagues (2017), which was then used for all segments.

5.2.11. Surrogate PLV

We assessed whether PLV and PDC values were significant compared to a baseline level. We created 1000 randomly paired infant adult dyads. Although we did not track the number of unique pairs from our n=50 datasets the total possible number of unique pairings was 1225. We then calculated PLV and PDC between these randomly paired dyads in exactly the same way, over the same frequency and channel combinations as in the real pair dyads.

5.2.12. Cluster correction for multiple comparisons

A cluster-based correction was applied to correct for multiple comparisons (we computed inter-brain PLV and PDC values for 32 x 17 x 480 combinations infant-adult data points) for the real pair infant adult inter-brain synchrony values. This was done by identifying clusters and removing those that contained lower than 75 contiguous channel x time x frequency points. The threshold of 75 was obtained using a non-parametric permutation procedure. Using the surrogate data generated above. We then cycled through this dataset 1000 times. On each iteration and for each channel, frequency and time point we randomly selected 1 random infant-adult pairing from the 1000 total pairings. We compared the PLV values for the selected pair against the PLV values of the full 1000 sample random pair dataset for each channel x time x frequency points. This was done by evaluating the selected pair PLV value's position relative to the 1000 sample random pair distribution using MATLAB's normcdf function. This generated a p value for each channel x frequency x time point for the selected pair PLV values. We then performed a second iteration of the data. For each iteration (and for 2D cluster correction for each channel separately) we thresholded the p values generated from the previous step at $p = .5$, significant PLV values were set to one and non-significant PLV values set to zero. Next, we used MATLAB's bwconncomp function to identify clusters of contiguously significant ($p < .05$) values in topographical time-frequency space. For each

iteration we stored the number of data points contained within the largest cluster identified. For some iterations there were obvious outliers in the sizes of the clusters identified. In these cases, outliers were removed using MATLAB's rmoutliers function. This gave us a distribution of 1000 cluster-sizes observed within the random pair PLV datasets. We then calculated the value for 95th percentile of this distribution. This gave us the threshold (75 data points). For clustering in 3-dimension space; clustering in topographical space was done by obtaining the interpolated (2D matrix) data image (with off-head points as NaN) of the topographical activity (scalp level) using EEGLAB's topoplot.m function. A topographical image was obtained for permutation (1000) and for each time-frequency point. This resulted in a 4D matrix (topographical images where the first 2 dimensions) for each permutation for which clusters where then identified. This approach is different to the approaches detailed in the previous chapters as we started with the brain activity instead of the behaviour.

5.3. Results

5.3.1. Behavioural synchrony

Firstly, we looked at whether infant's and caregiver's gaze and touch behaviours are coupled above chance during natural interactions. The results suggest that overall, infant's and adult's behaviours are significantly coupled during social interactions. Results of the cross-correlation analysis with various behaviours is reported in figure 1

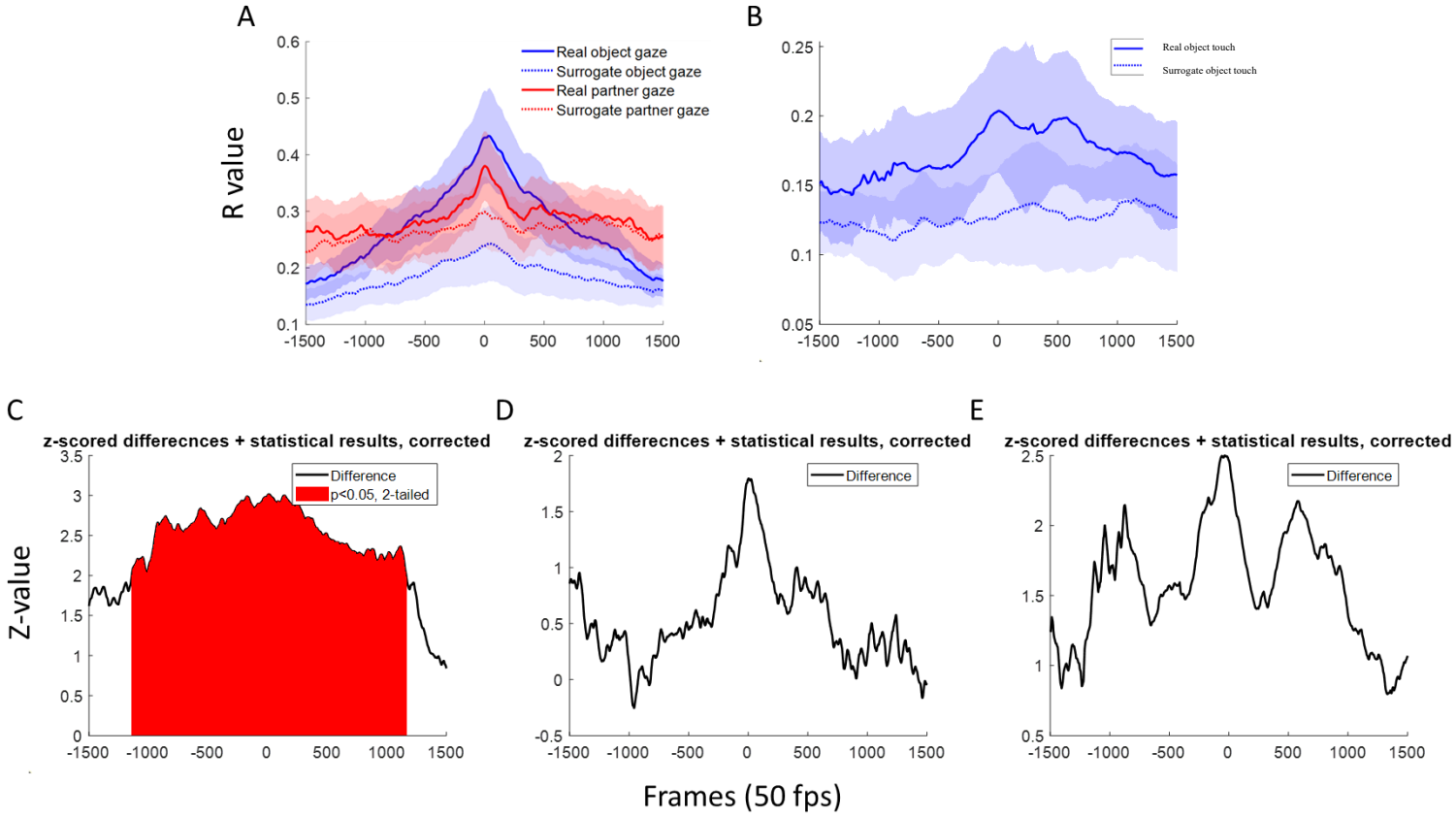


Figure 5.1. Shows the results of the cross-correlation analysis between parent infant gaze and touch behaviours compared to surrogate paired infant-adult data. A) shows the result of cross correlation analysis between object and partner looks for real and surrogate pair dyads. B) shows the result of cross correlation analysis between object touch for real and surrogate pair dyads. C) shows the results of the cluster-based permutation analysis comparing real vs surrogate paired object gaze data. D) shows the results of the cluster-based permutation analysis comparing real vs surrogate paired partner gaze data. E) shows the results of the cluster-based permutation analysis comparing real vs surrogate paired touch data. Red patches indicate areas where the real pair data was significantly more correlated than the surrogate pair data. Cross correlations were only significantly greater

than baseline for the object gaze. For all plots x axis shows cross correlation lags in frames, 1 lag = 1 frame (20ms).

Figure 5.1 shows the results of the cross-correlation analysis between parent infant gaze and touch behaviours compared to surrogate paired infant-adult dyads. We found that for all object, but not partner looks cross correlations between parent and infant dyads were significantly greater than baseline (cross correlation in surrogate pair data). The effects for touch were overall weaker and did not exceed baseline. Overall, the results of analysis 1 are consistent with previous research (Richardson et al., 2007) that show that during social interaction gaze behaviours are highly cross correlated. These results also extend these findings to other behavioural modalities e.g., touch.

5.3.2. Individual behaviours and infant EEG power cross correlations

Before turning to our main research question of whether parent infant inter-brain synchrony associates with behavioural coordination during social interaction we wanted to first investigate whether there were patterns of intra individual neural activity that associated with these behaviours. This could potentially be both an interesting theoretical point but also could confound the inter-brain synchrony analysis due to the impact of power on phase locking (e.g., see Marriott Haresign 2022a). We looked at associations between infant intra-brain power and individual looking and touch behaviours as well as associations between infant brain power and joint gaze and touch.

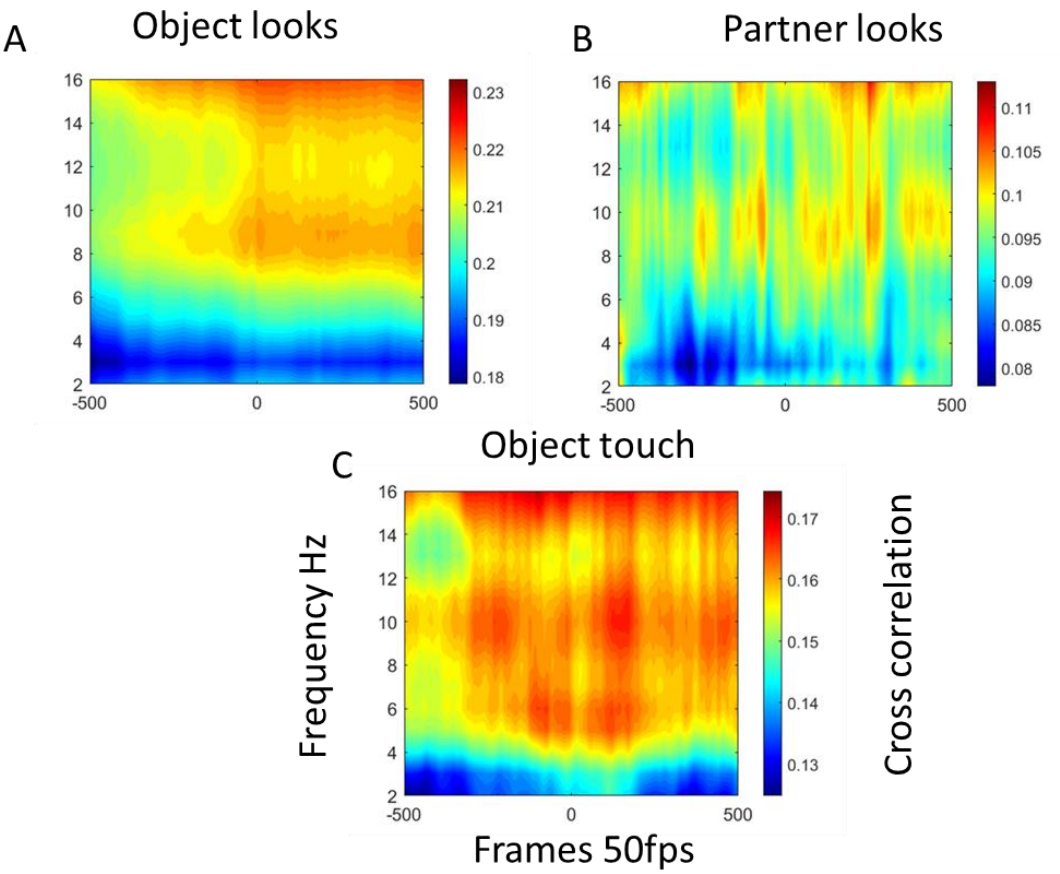


Figure 5.2. Shows the result of the cross-correlation analysis between infant EEG power (occipital electrodes) and individual gaze and touch behaviour. A) shows the result of cross correlation analysis between infant EEG power and object looks, averaged over all three objects. B) shows the result of cross correlation analysis between infant EEG power and partner looks. C) shows the result of cross correlation analysis between infant EEG power and object touch, averaged over all three objects.

5.3.3. Dyadic behaviours and infant EEG power cross correlations

We also wanted to examine associations between infant intra-brain power and joint gaze and touch. Figure 3 summarises the results of the cross-correlation analysis between infant EEG power and joint gaze and touch behaviours.

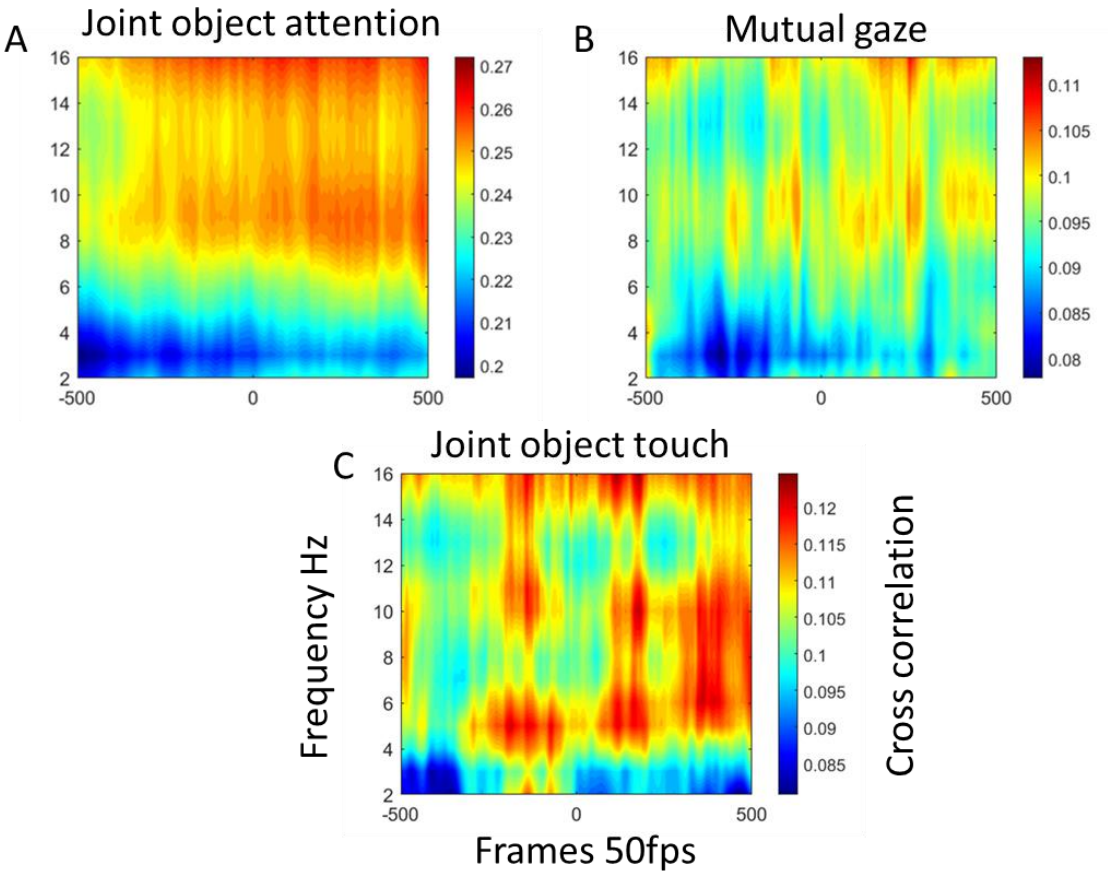


Figure 5.3. Shows the result of the cross-correlation analysis between infant EEG power (occipital electrodes) and joint gaze and joint touch behaviour with caregiver. A) shows the result of cross correlation analysis between infant EEG power and joint object attention, averaged over all three objects. B) shows the result of cross correlation analysis between

infant EEG power and mutual gaze. C) shows the result of cross correlation analysis between infant EEG power and joint object touch, averaged over all three objects.

Figures 5.2 and 5.3 show the results of the cross-correlation analysis between infant EEG power (averaged over occipital electrodes) and individual as well as joint gaze and touch behaviours. We compared this to surrogate paired data (random EEG times series paired with random behavioural times series). We used a non-parametric cluster-based permutation analysis to identify which frequencies and time points were significantly greater in the real paired data. The results of the permutation analysis revealed that no associations between infant EEG power and individual or joint gaze and touch behaviours were significantly greater than the surrogate data.

Overall, the results of the analysis presented in sections 5.3.2 and 5.3.3 are consistent with previous research (Wass et al., 2018) that show that associations between infant EEG power and gaze behaviour during joint play are weak. However, these results also extend previous work in three main ways. A) by looking at whether these association are significantly greater than surrogate which was not previously done. B) by looking at associations between infant EEG power and joint attentional processing with the caregiver. C) by looking at touch as an additional behavioural modality.

5.3.4. Inter-brain synchrony

To investigate the relationship between inter-brain synchrony and behavioural coordination, we first computed mean PLV values across the entire interaction. We looked at whether PLV values were significantly greater than baseline, and then whether significant clusters of inter-

brain synchrony associated with behavioural coordination. Based on a number of previous findings which have suggested that inter-brain synchrony is increased during moments of behavioural coordination we expected to observe significant associations between these variables.

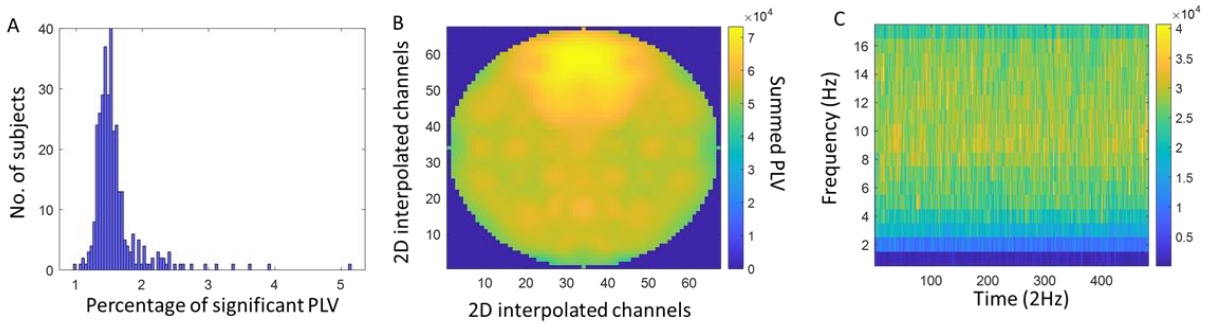


Figure 5.4. Illustration of 3D clustering procedure. A) shows distributions of percentage of above surrogate PLV across the entire interaction. B) shows summed topographical distribution of above surrogate PLV values over all participants. B) shows marginal time-frequency distribution of above surrogate PLV values, summed over all participants

Figure 5.4 shows the results of analysis of inter-brain synchrony. We first tested whether PLV values significantly exceeded surrogate data. Looking across the entire interaction the results of the permutation analysis did indicate some significant clusters of PLV; on average this accounted 1.6% (SD = 1.9) of the total possible PLV values (36630240) that we searched over.

5.3.4. Inter-brain synchrony cross correlations with dyadic behaviour

We next wanted to examine whether the above chance PLV values were significantly associated with behaviour compared to chance levels generated from cross correlations with surrogate datasets. Again, here we looked at cross correlations between above chance PLV and gaze and touch behaviours. The results of this analysis are summarised in figure 5.

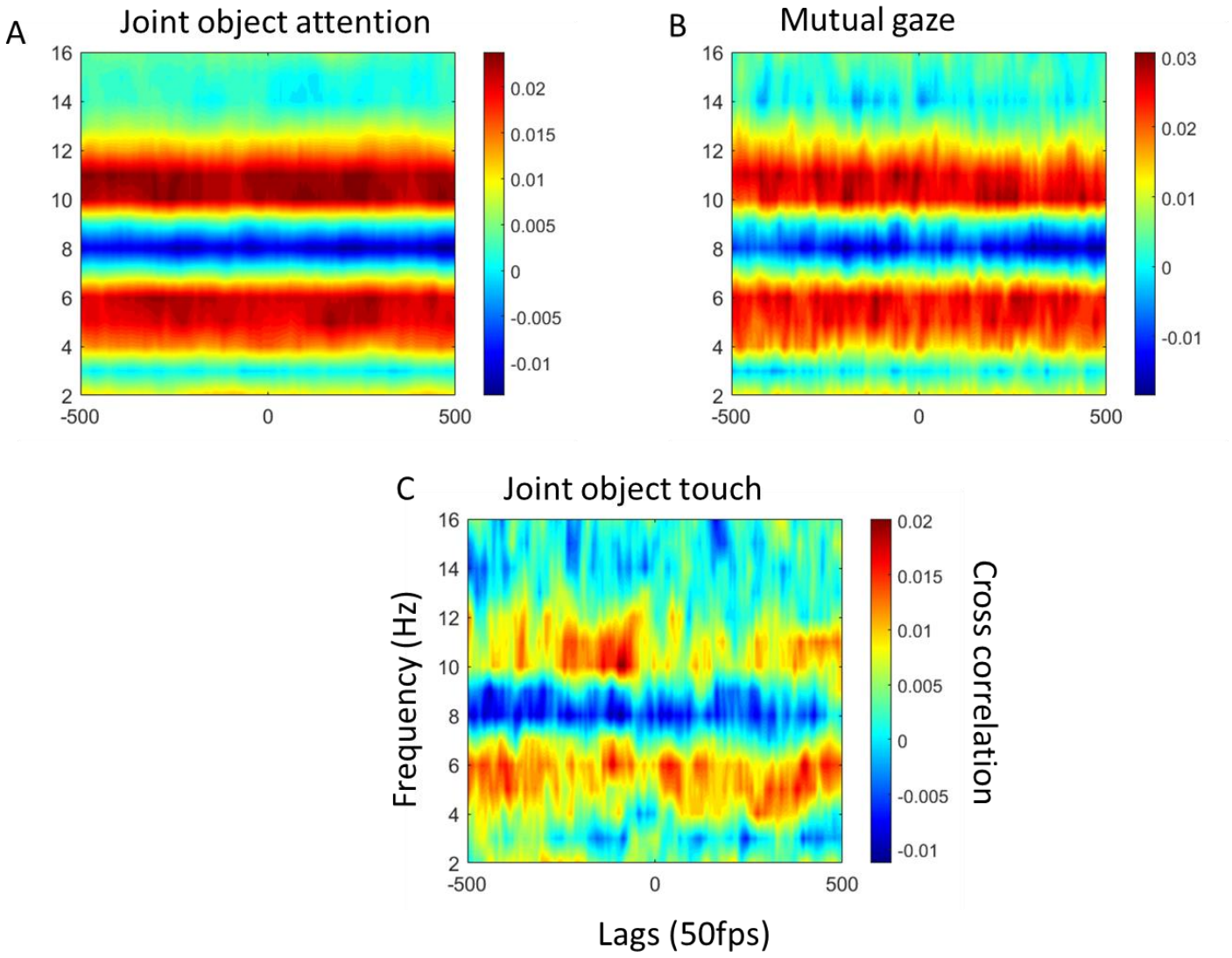


Figure 5.5. Shows the result of the cross-correlation analysis between parent-infant PLV and power (averaged over all electrodes) and joint gaze and joint touch behaviour with caregiver. A) shows the result of cross correlation analysis between parent-infant PLV and

joint object attention, averaged over all three objects. B) shows the result of cross correlation analysis between parent-infant PLV and mutual gaze. C) shows the result of cross correlation analysis between parent-infant PLV and joint object touch, averaged over all three objects.

Figure 5.5 shows the results of the cross-correlation analysis between PLV, and parent-infant joint gaze and touch behaviours compared to surrogate paired infant-adult PLV values. We used a non-parametric cluster-based permutation analysis to identify which frequencies and time points were significantly greater in the real paired data. The results of the permutation analysis revealed that no associations between PLV and joint gaze and touch behaviours were significantly greater than the surrogate data. Overall, the results of analysis presented in section 3.4 suggest that there are patterns of above chance inter-brain synchrony that arise during social interactions and that these do associate with parent-infant behavioural coordination. However, these association were not greater than associations that arose by chance.

5.3.4. Inter-brain synchrony cross correlations with intra-brain power

If behavioral coordination is not driving above chance inter-brain synchrony what other factors are? One factor that might impact inter-brain synchrony values in surrogate (and real data) is intra-brain power. It is known that there is a relationship between EEG power and phase. For example, Muthukumaraswamy and colleagues (2011) showed that increases in power can lower error in phase estimation and give the appearance of heightened phase locking (see also, Burgess, 2013). Separating increases in power from genuine increases in phase locking is difficult and continually debated (e.g., Sauseng et al., 2007). Therefore, one

factor that might impact inter-brain synchrony values in surrogate (and real data) is intra-brain power. To explore this possibility, we examined the relationships between infant and adult EEG power and inter-brain synchrony (PLV).

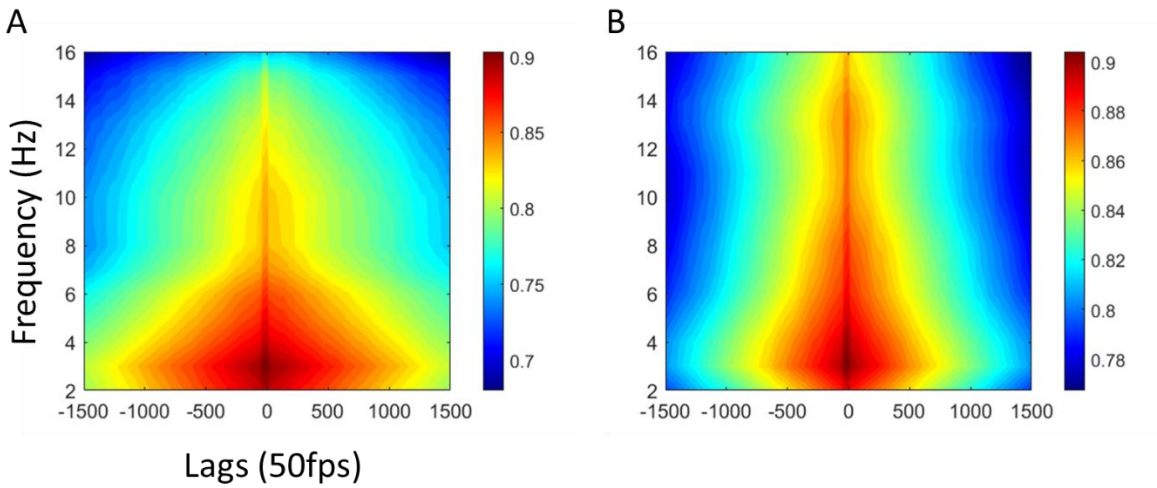


Figure 5.6. Shows the result of the cross-correlation analysis between intra-brain power and inter-brain PLV. A) shows the result of cross correlation analysis between infant EEG power and PLV (averaged over all electrodes). B) shows the result of cross correlation analysis between adult EEG power and PLV (averaged over all electrodes). For both infant and adult data all time-frequency points were significantly greater than surrogate data.

5.4. Discussion

We took dual EEG recordings from parents and infants whilst they engaged in naturalistic free-flowing social interactions. We computed inter-brain synchrony following cleaning and

analysis procedures specially designed for naturalistic dual EEG data (Haresign et al., 2021; 2022). While many studies of inter-brain synchrony have explored how these patterns are distributed across two dimensions (e.g., channels x frequency) no studies have analysed the entire data space (i.e., in topographical, time and frequency space). Given that this field is still relatively new and there is a deficit of empirical research to generate new and testable hypothesis such an approach would have the obvious advantage of not having to relying on predefined/ researcher chosen locations, frequencies and time points. Through this we wanted to consider two mains sets of research questions:

Behavioural synchrony

For our first set of research questions, we examined whether infant's and caregiver's behaviours are coupled above chance during natural interactions. The results of the cross-correlation analysis suggested that overall, infant's and caregiver's object gaze and touch behaviours towards objects were significantly coupled. These results are in line with previous research (e.g., Richardson et al., 2007) that shows that gaze patterns between interacting individuals are highly coupled during social interactions.

Intra-brain analysis

Before turning to our main research questions of whether parent infant inter-brain synchrony associates with behavioural coordination during social interaction we wanted to first investigate whether there were patterns of intra individual neural activity that associated with these behaviours. This could potentially be both an interesting theoretical point but also could confound the inter-brain synchrony analysis due to the impact of power on phase locking (e.g., see Marriott Haresign 2022a). To investigate this, we looked at associations between

infant intra-brain power and individual looking and touch behaviours as well as associations between infant brain power and joint gaze and touch. For each data for the real pair dyads was compared against surrogate data generated through random permutation of dyad membership (e.g., random EEG times series paired with random behavioural time series) using cluster-based permutation analyses. The results of these tests indicated no significantly above chance relationships between infant intra-brain power and any of the behavioural measures we investigated; solo and joint looking at objects, partners as well as solo and joint touching of objects. These results are unsurprising given that previous research found only very weak correlations between infant EEG power and object gaze during joint tabletop interactions (e.g., Wass et al., 2018). Although Wass and colleagues did not compare cross correlations against baseline data, the results of our study suggest that for joint play with caregivers' infant's EEG power does not associate significantly with onsets of object gaze or touch behaviours.

Evaluating 3-dimensional inter-brain synchrony across the entire interaction

Our main set of research questions explored whether, across the entire interaction significant patters of adult-infant inter-brain synchrony arose and whether significant patterns of inter-brain synchrony associated with moments of parent-infant behavioural coordination. Firstly, significant inter-brain synchrony was assessed by comparison to surrogate data generated through random permutation of dyad membership. This process was done individually for each participants. The results of the cluster-based permutation analysis indicated that for every participant the were significant (3D) patterns of inter-brain synchrony (PLV) that arose during the interaction. However, these above chance PLV values within the data were very subtle accounting for on average only 1-2% of the total number of PLV values. These findings are consistent with the growing body of positive evidence of the presence of inter-

brain synchrony during social interactions. However, where previous studies have focused their investigations on one or two dimensions of the data (e.g., frequency, frequency x topography), we took an entirely data driven approach, searching statistically significant patterns of inter-brain synchrony across all dimensions (time, space and frequency) simultaneously. Through application of this approach, we found that whilst statistically significant patterns of parent-infant inter-brain synchrony were increasable subtle accounting for only a very small number of the total possible PLV values.

Whilst our study shows the efficacy of using data driven methods to identify patterns of inter-brain synchrony and attempt to understand the mechanisms that contribute to it, it is not without limitation. Firstly, the computational time and intensity of the analytical approach that we have followed with this paper is vast. Typically studies that use 3+ dimensional cluster-based permutation analyses do so with event locked data, and don't often search the entire data space. Researchers should weigh the amount of computational time required to obtain 2D spatial topographical maps, build distributions of 4D data and identify clusters within this, against what implications from the data can be drawn through application of this approach. Additionally, whilst our approach is well suited for exploratory data analysis as the field of inter-brain synchrony becomes more developed and research questions becomes more theoretically driven approaches like these that search over the entire data space may become less important.

Inter-brain synchrony and behavioural coordination

The main aim of this study was to explore the relationship between inter-brain synchrony and behavioural coordination. Although very little is known about the mechanisms that give rise to inter-brain synchrony most of the current theoretical accounts confer that inter-brain

synchrony during social interactions must be embedded within interpersonal behavioural coordination (Dumas et al., 2011, Burgess 2013, Wass et al., 2020, Hamilton, 2021, Novembre et al., 2021). Without tangible links to real world events/ processes it is very difficult to understand (and largely speculative) how two brains might become synchronised, and further this can muddy the water between different sources of inter-personal neural associations (Holroyd, 2022). We looked at associations between inter-brain synchrony and moments of spontaneous parent-infant behavioural coordination during free-flowing naturalistic social interaction. We reasoned that if inter-brain synchrony is indeed grounded in moments of behavioural coordination, we should observe a temporally sensitive relationship that is greater that what might be observed by chance – using surrogate data generated through permutation of dyad membership. To explore this, we extracted moments of object gaze and touch coordination during parent-infant joint play interactions. We then looked at associations between moments of behavioural coordination and above chance patterns of inter-brain synchrony, again comparing this to surrogate data generated through random permutation of dyad membership (random PLV time series with random behavioural time series). The results of this analysis revealed that whilst there were strong associations between above chance PLV and behavioural coordination for both gaze and touch modalities, these associations were not greater than the correlations values that would have been found in the data by chance (based on comparison to surrogate cross correlation values). Overall suggesting that inter-brain synchrony was not grounded within the real time coordination of behaviours between interaction infants and parents.

The associations we observed between inter-brain PLV, and intra-brain power are highly similar to those previously observed between intra-brain PLV and intra-brain power (e.g.,

Muthukumaraswamy et al., 2011), further emphasising the importance of interpreting concomitant changes intra-brain EEG power and inter-brain synchrony cautiously.

5.5. Conclusion

In this study we analysed patterns of inter-brain synchrony in parent-infant dyads during naturalistic social interactions (re-analysis of a N=55 parent-infant, approx. 12m old, dual EEG dataset). We took an exploratory data driven approach, investigating patterns of above chance inter-brain synchrony in topographical, frequency and temporal space using nonparametric statistics and cluster correction across the three dimensions. Our aims were to first identify where above chance patterns of inter-brain synchrony were arising during social interaction and second to explore whether these were grounded in moments of behavioural coordination. The result of our analysis suggests that consistent with a growing body of positive evidence parent-infant inter-brain synchrony did arise during free-flowing social interactions, although these patterns were significant only a very small percentage of the time (1-2%). And whilst we did find evidence to suggest that significant patterns of inter-brain synchrony associated with behavioural coordination in a time sensitive way (e.g., peak in cross correlation at time 0), comparison to surrogate data revealed that these associations were not greater than what arose by chance based on the inherent variation within the data. Overall, the result of our analysis emphasises the importance of appropriate statistical approaches that control for the amount of inter-brain synchrony that might arise in the data by chance, something that, as noted recently (Holroyd, 2022), not all investigations of inter-brain synchrony do. Crucially our results failed to provide support for theories of inter-brain synchrony that are grounded in moments of behavioural coordination (Dumas et al., 2011, Burgess 2013, Wass et al., 2020, Hamilton, 2021, Novembre et al., 2021). This further

emphasises the need to focus on what mechanisms that might give rise the inter-brain synchrony, e.g., a) what changes and to which aspects of the EEG signal are necessary to facilitate inter-brain synchrony (see Haresign et al., 2022) and (b) what are the known behavioural, cognitive processes that might trigger such changes. Through this we can develop a more in depth understanding of inter-brain synchrony rather than the disparate links between we currently understand about social behaviour and inter-brain synchrony.

Chapter 6 – Discussion

Despite a growing prevalence in empirical evidence linking behavioural coordination during social interaction with inter-brain synchrony, these correlates have been routinely studied using highly structured tasks and analysed as a time invariant phenomenon. The present thesis explored the methodological challenges and benefits of recording dual EEG activity from parents and 1 year old infants whilst they engaged in naturalistic/ unstructured social interactions. To investigate inter-brain correlates of parent-infant social interactions, across multiple analyses, parent-infant dual EEG data was aligned with moments of naturally occurring behavioural coordination extracted from high resolution (50 fps) video recordings and patterns of intra and inter-brain activity were examined. The first part of this work focuses on the unique challenges associated with measuring/ analysing intra and inter-brain activity within this dataset. Having validated our methodology, the second part of this thesis focuses on the mechanisms that give rise to inter-brain synchrony during early social interactions. Here we first examine fine-grained changes in parent-infant inter-brain synchrony time locked to naturally occurring mutual gaze onsets. Second, we examine

associations between inter-brain synchrony and behavioural coordination across multiple modalities.

Eye movement artifacts in naturalistic EEG data – review

To date social neuroscientists investigating infants' social brain functioning have routinely used experimenter-controlled paradigms, in which information is presented to infants via a screen, contingent upon their sustained fixation. In the real world however, we know that information is not presented to us in this way, but rather that humans constantly, visually 'forage' for information from their environment, as indexed by the fact that humans tend to make a saccadic eye movement approximately every 300-400ms. Research designs that restrict natural vision to fixation, likely due to reasons associated with eye movement artifact (discussed in more detail below), crucially assume two main things that hinder a more in depth understanding of social brain function. Firstly, that the muscle and neural activity associated with the ocular movement itself is not relevant for cognition and second that neural responses to stimuli are the same when information is presented contingent on fixation vs contingent upon a saccade.

Recently, there has been a drive towards the use of more naturalistic paradigms in developmental EEG research (Risko et al., 2016; Wass et al., 2020; Holleman et al., 2020). However, naturalistic EEG recordings provide additional analytical challenges over traditional screen-based tasks. For example, in traditional screen-based, event-related tasks in which the child is passively exposed to a set of stimuli, artifacts are more randomly distributed with respect to the simulation. Removal of sections containing significant artifact

can in this context be potentially beneficial, as visual experience during these sections might also be different (e.g., at its simplest the child might be fussing and not be attending to the image on the screen). However, in naturalistic paradigms, removal of whole sections of data is particularly problematic because data segments contaminated by artifact often covary with cognitive/ attentional processes of interest. Specifically, in naturalistic paradigms, the 'simulation' can be child-controlled (e.g., the child turning to the parent in a naturalistic interaction), and so artifacts are more likely to be time-locked to neural signals of interest. The removal of artifact is thus likely to also affect the analysis of neural signals.

As no methods existed that were tailored for correcting time locked artifacts in infant EEG, it was necessary to first develop one – which was the basis of the work that contributed to original chapter 5. To do this we adapted one of the most popular systems for adult ICA component classification (artifact of not) for use with infant EEG data. We then tested the classifiers' ability to clean ERP data time locked to moments when the infant looked towards their caregiver. We found that for our infant EEG data the adapted classifier removed more activity presumed to be artifact and retained more activity presumed to be neural than the original classifier. Whilst we could not conclude from the results (as few studies can) that our system removed all artifact, our adapted classifier provides a tool for automatic artifact removal for infant EEG data, application of which at the very least will be necessary for visualisation of neural signals which are often orders or magnitude smaller than artifactual signals.

The contributions of this work were three-fold. Firstly, there was a clear need within the community for tools that were specifically designed for infant EEG data (almost all comparable systems are built around and designed for adult EEG data), as indexed by the

considerable attention this system has already received (15 citations in the first six months of publication). Secondly, consistent with previous research which suggests that averaged EEG activity time locked to the offset of saccades (lambda activity) and ERPs elicited during sustained fixation (visually evoked potentials) share common neural generators (Kazai and Yagi, 2003). The ERPs that we demonstrated in this work showed remarkable similarity to infant ERPs previously obtained using screen-based paradigms (e.g., Guy et al., 2016; 2018; Peykarjou & Hoehl, 2013; Xie & Richards, 2016). That our ERPs showed a close visual correspondence with ERPs observed in traditional screen-based ERP paradigms suggests that the use of a novel, naturalistic paradigm to explore time locked changes in infant's neural sensitivity was valid, and we had developed a well performing tool that was capable of addressing the unique artifactual challenges associated with this study design.

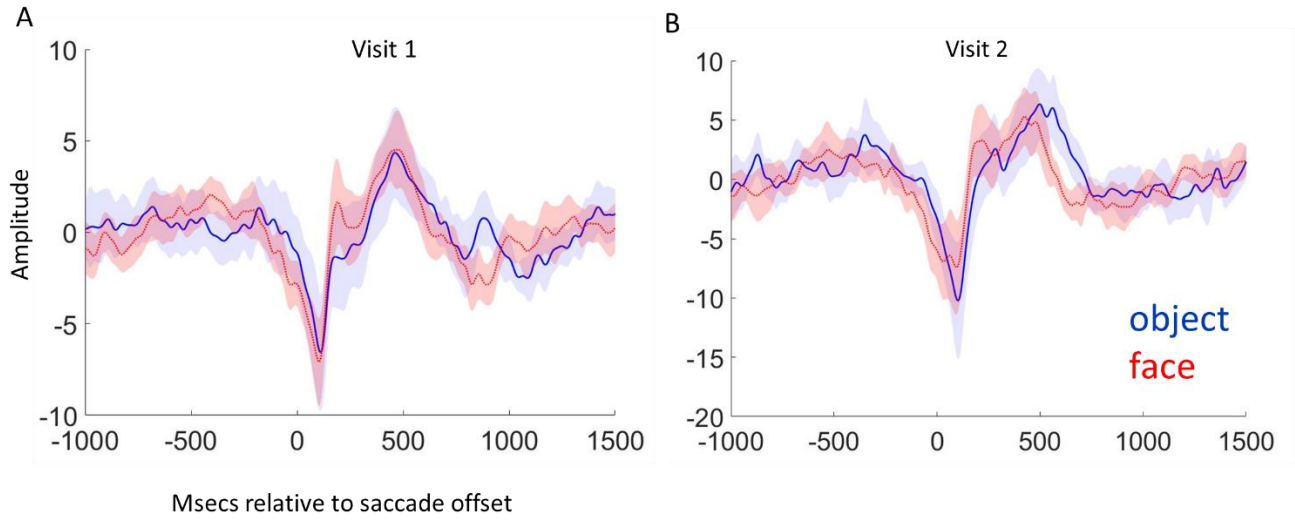


Figure 6.1. Event-related potentials time-locked to naturally occurring partner (face) gaze and object gaze onsets. A and B both show infant occipital ERPs relative to onsets of infant

looks to their partners face vs infant looks to objects. Shaded areas indicate 95% confidence intervals, thicker lines indicate grand average waveforms. Data for A was taken from visit 1 of the 8-week research programme and data for B was taken from visit 2. In both sets of data, we replicated previously obtained results using screen-based paradigms (e.g., Guy et al., 2016; 2018; Peykarjou & Hoehl, 2013; Xie & Richards, 2016), i.e., a greater amplitude N290 ERP component for faces vs objects.

Future directions

Developing more robust metrics of pre-processing performance. Beyond making this tool more useful the field as a whole would benefit from more quantifiable techniques to assess whether or not an EEG pre-processing procedure is objectively improving the quality of the data (e.g., Delorme, 2022). The difficulty in this is determining the 'ground truth' upon which to make comparisons. Research like that which forms the basis of chapter 5 of the present thesis routinely assess performance against data that is manually scored for artifact by one or two experienced researchers. However, this is far from perfect, increasing the number of researchers that independently score the data and taking averages from this would improve the reliability of 'ground truth' data. Tools like the one presented in chapter 5 are useful for this purpose as they publish a dataset that has been manually scored. However, ultimately as pre-processing procedures are largely study specific the field will need to develop measures that are capable of assessing the performance of multiple pre-processing steps on any given dataset.

Understand cleaning procedures are limited and improving how we report artifact. Studies like the one presented in chapter 5 highlight the fact that data cleaning procedures are limited and are not capable of completely removing activity associated with artifact from EEG data. As social neuroscience continues to move towards more naturalistic study designs it will become increasingly important for researchers to examine the contribution of artifact to their data. For example, even very small eye movements can produce artifacts in the EEG capable of creating between condition differences (e.g., Dimigen et al., 2009). Therefore, at a basic level it will be important for research using naturalistic study designs to examine if there are behavioural (e.g., amount or size of eye movements) differences between groups/ conditions that could contribute to differences in the EEG.

Understanding the neural correlates of active vs passive visual processing. Although as we noted (see Figure 4.3) the ERPs that we observed share a close correspondence with ERPs observed in traditional screen-based ERP paradigms one fundamental question we were not able to address in this work is what was driving the differences between our ERPs and previous screen-based ERPs. Most notably we observed a large negative potential, peaking occipitally around 100ms following the end of an eye movement. This can be clearly seen in figure 3. This feature of the occipital ERP waveform is not present in comparable screen-based studies that have explored infant's neural responses to images of faces (e.g., Farroni et al., 2002). One possibility is that the early negativity we observed could be the result of residual artifiact – generated from dipole activity of the eye movement itself. This is indeed possible as one limitation of this work was that we were not able to completely remove the activity we assumed to be related to the eye movement artifact – something that is consistent even with studies that use much cleaner adult EEG data (e.g., Plöchl et al., 2012).

Another possibility is that it reflects genuine neural activity driven by a process involved in active vision (e.g., visual processing accompanied by a voluntary eye-movement), that was not present in previous passive, screen-based tasks. Which might only be captured by studying neural responses using naturalistic study designs. Unfortunately, with this current dataset we were unable to fully disentangle these two possibilities. In current experiments and follow up work we aim to explore this in more detail using co-registered EEG and eye-tracking, in which we contrast conditions in which information (visual target) is presented contingent upon a saccade vs when no information is presented but a saccade is still elicited and when information is presented contingent on fixation. Thus, we can further investigate infant's neural correlates of eye movements and compare neural responses to information presented contingent upon eye movements vs fixations.

Inter-brain synchrony during naturalistic social interactions

The second part of this thesis explored the challenges and benefits associated with measuring parent-infant inter-brain synchrony during naturalistic social interactions. As highlighted in the introduction one thing that has limited our understanding of inter-brain synchrony so far is that the field has primarily adopted approaches that measure inter-brain synchrony as a time invariant property – meaning that these studies typically measure how inter-brain synchrony varies topographically or across frequencies but not through time. This approach crucially obscures important information about how inter-brain synchrony develops during and supports social interaction. Early parent-infant social interactions provide the ideal setting to study these mechanisms as temporal coordination during these social interactions

has been shown to be integral in supporting infants' later language outcomes and cognitive development.

The present thesis aimed to take the first steps in addressing this deficit. The primary goal was to test the hypothesis laid out by previous work by Drs (and supervisors of the current thesis) Wass and Leong (*Speaker gaze increases information coupling between infant and adult brains*, Leong et al., 2017) which suggested that inter-brain synchrony was stronger bidirectionally during mutual relative to non-mutual gaze. The present thesis aimed to reproduce the mutual gaze - inter-brain synchrony effect in naturalistic parent-infant recordings, and further, to investigate whether the effects of mutual gaze on inter-brain synchrony were present around onsets of mutual gaze. Subsequently, we wanted to assess how inter-brain synchrony arises around mutual gaze onsets by exploring patterns of intra and inter-brain activity.

To do this we extracted concurrent dual EEG activity around naturally occurring gaze onsets during infant-caregiver social interactions. These experiments generated several main insights. First the methodological problems associated with co-occurring eye movement artifacts that were discussed in the previous section are manageable to an extent that permits the replication of established ERP effects (see figure 4.3). Second the effects of mutual gaze onsets are strongest at the intra-brain level, in the 'sender' but not the 'receiver' of mutual gaze. Third, despite a growing prevalence of theories within the wider research field suggesting that increased inter-brain synchrony associates with mutual gaze, here in the current sample of parent-infant (approx. 12M) dual EEG recordings we did not find that mutual gaze (onsets) was associated with increases in inter-brain synchrony over surrogate data or relative to moments of non-mutual gaze.

ERP effects. A first basic question when exploring neural sensitivity to mutual gaze onsets in natural settings is can previous findings using screen-based simulacra be replicated. In 2002 Farroni and colleagues', reported evidence from their study which suggested that 4-month-old infants show different neural responses contingent on whether someone is looking directly at them or not. Evidence for enhanced ERP amplitudes for faces with direct vs indirect gaze has already been extended to live contexts (Pönkanen et al., 2011a, but not to unstructured, free-flowing, social settings. However, as we noted ERP findings on the effects of mutual gaze have not replicated well in developmental research (e.g., Elsabbagh et al., 2009), or in research with adults (e.g., Watanabe et al., 2001; Taylor et al., 2001b; Watanabe et al., 2002; Itier et al., 2007; Conty et al., 2007; Pönkanen et al., 2011a). In chapter 4 we extracted dual EEG activity around naturally occurring gaze onsets during infant-caregiver social interactions in N=55 dyads (mean age 12 months). We differentiated between two types of gaze onset, depending on each partners' role. 'Sender' gaze onsets were defined at a time when either the adult or the infant made a gaze shift towards their partner at a time when their partner was either already looking at them (mutual) or not looking at them (non-mutual). 'Receiver' gaze onsets were defined at a time when their partner made a gaze shift towards them at a time when either the adult or the infant was already looking at their partner (mutual) or not (non-mutual). Contrary to our hypothesis based on previous findings (e.g., Farroni et al., 2002) we did not find evidence for increased ERP amplitudes around mutual vs non-mutual gaze onsets.

One possible explanation for the inconsistencies between previous screen-based research and the present study is simply it is just a result of increased artifact through the use of a naturalistic paradigm. However, we note that in first-author publications 1 and 3 our ERPs

show a close visual correspondence with ERPs observed in traditional ERP paradigms; the overall measures of EEG data quality we reported show good quality data (see section 8 of supporting materials for publication 3); we did replicate the findings from screen-based ERP research that infants show enhanced ERPs to images of faces vs objects (Guy et al., 2016; 2018; Peykarjou & Hoehl, 2013; Xie & Richards, 2016) (see section 1 of supporting materials for publication 3); We observed statistically greater occipital ERP amplitudes for faces vs objects for the N290 component, but not for P1 or P400 components (ERP components typically routinely examined under this type of experimental paradigm/ stimuli.

Mutual gaze and inter-brain synchrony. Secondly, in the present thesis we wanted to examine neural sensitivity to mutual gaze onsets at the dyadic level – e.g., do patterns of inter-brain synchrony between parents and infants associate with these processes. We wanted to reproduce the mutual gaze inter-brain synchrony effect, found in previous studies (Leong et al., 2017), using naturalistic parent-infant recordings. We also wanted to investigate whether the effects of mutual gaze on inter-brain synchrony were present around onsets of mutual gaze. The results of chapter 7 (publication 3) were inconsistent with previous studies that observed greater inter-brain synchrony during continuous (i.e., not relative to specific behaviours/ events within the interaction, but rather looking across all moments of a given behaviour during social interaction) moments of mutual vs non-mutual gaze. For example, Leong and colleagues (2017) found increased inter-brain synchrony using PDC, in Theta and Alpha, over C3 and C4 electrodes in $N = 29$ *8*-month-olds (Leong et al., 2017). In the present study, we measured PDC and PLV across the same frequencies and electrodes in $N = 55$ *12*-month-olds.

Although we followed the same analytical techniques as Leong et al., 2017, we used different pre-processing techniques and a different (less structured, more naturalistic) paradigm, which

could explain why our results differ. Firstly, the previous study featured an unfamiliar live adult singing nursery rhymes to an infant. Our present study, in contrast, featured primary caregivers interacting freely with their infant, using a puppet that they placed on their hand. Infant's sensitivity to novel interaction partners is well documented (Bushnell et al., 1989; de Haan & Nelson, 1997; 1999; Barry-Anwar et al., 2016; Hoehl et al., 2012). Therefore, one explanation for the positive effects of gaze type on inter-brain synchrony in the previous study could be due to the saliency of mutual gaze in the presence of an unfamiliar adult. To investigate this further we performed the same analyses with data collected from infant-adult dyads. Here the infants interacted with an unfamiliar adult (one of two research assistants). The results of these analyses are reported in full in the supplementary materials for chapter 7, but summarised here, we found consistent with our main analysis of infant-caregiver dyads that inter-brain synchrony was not above chance around mutual gaze onsets and did not differ between mutual and non-mutual gaze onsets. Further we found that phase resetting around mutual gaze onsets was strongest for infant and adult sender compared with receiver gaze onsets.

Second, in the previous study (Leong et al., 2017), adults continuously sung nursery rhymes to the infants during the interactions, whereas in the present study (chapter 7, first author publication 3) they talked normally. As sung nursery rhymes are highly periodic (Suppanen et al., 2019) and evidence suggests that infant's neural activity entrains to the temporal structure of these songs (Leong et al., 2017a; Attaheri et al., 2022), it could be that the regularity of the nursery rhymes introduced an external periodic stimulus into the environment that was driving the inter-brain entrainment (e.g., Perez et al., 2017). Here, mutual gaze might only enhance or maintain synchrony that is already established, by facilitating shared attention and therefore upregulating attention-enhanced neural synchrony.

It will be important for future research to examine inter-brain synchrony in a variety of settings, ranging from very unstructured settings such as those used in the present study to more structured settings in which there are environmental stimuli with more regular and predictable inputs.

Additionally, to address the possibility that differences in pre-processing procedures could explain why we failed to replicate previous findings we re-cleaned our data following the pre-processing procedures outlined in Leong and colleagues' (2017) study. We found that re-cleaning the data following the procedures of Leong and colleagues had no impact on the significance of any of the results of the main paper (publication 3), ruling out the possibility that different pre-processing procedures might be the cause of the discrepancy. Overall, these inconsistencies highlight the likely context-specific and localised nature of inter-brain synchrony, and further emphasise the importance of replication and standard data quality measures (Luck, 2021) when studying inter-brain dynamics (Holroyd, 2022).

Phase resetting around gaze onsets. In chapter 4, we also explored event-locked intra and inter-brain neural responses associated with mutual gaze onsets. Through this, we aimed to test our previously published hypothesis that concomitant phase resetting in the sender and the receiver's brain at the onset of gaze may drive inter-brain synchrony (Leong et al., 2017; Wass et al., 2020). Overall, the results of our event-locked analyses are inconsistent with this idea. Contrary to our hypothesis, inter-brain synchrony did not significantly exceed baseline values for sender/ receiver mutual or non-mutual gaze onsets and was not significantly different between sender or receiver mutual vs non-mutual gaze onsets.

Further, whilst we found that sender but not receiver mutual and non-mutual gaze onsets led to significant increases in ITC and amplitude (ERPs) over baseline, we did not find significant differences between sender or receiver mutual vs non-mutual gaze onsets. We did, however, find evidence for increases in ITC relative to sender mutual and non-mutual gaze onsets, but it is difficult to conclude that this represents phase resetting of neural oscillations. It could also be that changes in event locked amplitude/ power created the artifactual appearance of phase synchrony (Muthukumaraswamy et al., 2011) – a fact that the close correspondences we observed between ITC and event-locked changes in amplitude/power (see supporting materials in chapter 7) would appear to support.

One possible driver of the sender neural responses could be residual eye movement artifact in our data. In supplementary analyses we compared time-frequency power over frontal and occipital electrodes before and after ICA cleaning and report that ICA cleaning removed most, but not all, of the assumed artifactual activity associated with the eye movement- a conclusion consistent with our previous research (Haresign et al., 2021). This analysis also allowed us to identify that these artifacts are transient (~100ms) and therefore only impacted the initial part of the ERP waveform. After the initial ~+150-200ms we observed ERP components that look very similar to ERPs observed in traditional screen-based tasks with clear P1, N290 and P400 components. For added safety, however, our main analyses were based on comparing sections of the data that are both identically time-locked to saccades, and therefore contain an identical amount of eye movement artifact.

Overall, then, the results of publication 3 challenge the theory that phase resetting around key communicative signals such as mutual gaze is a mechanism through which inter-brain synchrony is achieved. Assuming that inter-brain synchrony according to more recent

frameworks (Holroyd, 2022) is associated with mutual gaze. This points to the importance of other potential drivers of inter-brain synchrony, that future work should investigate in more detail – such as correlated changes in amplitude/ power or changes in oscillatory frequency independent of phase resetting (see Haresign et al., 2022 for a detailed discussion), and other more periodic behaviours (e.g., speech; Leong et al., 2017a; Attaheri et al., 2022).

Future directions

Development of statistical procedures for analyses of inter-brain synchrony. A key issue in the study of inter-brain synchrony is how to isolate the distinct contribution of the real time social interaction from more general environmental features/ differences in cognitive engagement associated with live social interactions. This is because as several authors have noted common neural entrainment to exogenous/ environmental stimuli can give the appearance of inter-brain synchrony (Burgess 2013; Holroyd 2022). For example, this could happen if two participants simultaneously experience the same stimuli such as one person speaking. Here, the speech signal could drive entrainment (i.e., synchronised oscillatory activity) between the speech and neural oscillations in both participants. Although this could give the appearance of increased inter-brain synchrony, similar patterns of inter-brain synchrony could also be observed if both participants neural responses to the speech signal were recorded and analysed in isolation (although see Pérez et al., 2017).

To address this issue of common neural entrainment many study designs include conditions in which participants perform similar behaviours to those in the live social condition and in the same environment, but without interacting directly (for an example see., Wass et al., 2018, Reindl et al., 2022). Additionally, many studies control for the issue of common neural

entrainment using statistical analyses in which the observed/ real inter-brain synchrony values are compared to 'surrogate' inter-brain synchrony values. Typically, surrogate datasets are generated by randomly permuting some aspect of the data (e.g., the time dimension or which members of the dyad are paired) and then recomputing inter-brain synchrony. Although these analyses offer increased interpretational value over not performing any additional statistics, recently criticisms have been raised over the efficacy of current statistical approaches for group level analysis of inter-brain synchrony (Holroyd, 2022). These criticisms are primarily to do with; a) how the surrogate data are generated and b) how the observed inter-brain synchrony values are compared to the surrogate values. Although these criticisms are based on more fundamental consequences of permutations statistics, this has yet to be empirically tested for inter-brain synchrony analyses. Future research could explore these criticisms using a combination of empirical and simulated data. More specifically – what is the likelihood of generating false positives using different permutation approaches. This work would be useful in guiding future investigations of inter-brain synchrony in how best to approach group level analyses.

Building our understanding of the neurocognitive mechanisms that underlie real time social behaviours. As has been recently highlighted (e.g., Cañigueral et al., 2022) much can be learned from questioning how we behave differently in the presence of a social partner and how our brains support this. The present thesis studied free-flowing naturalistic social interactions to investigate the mechanisms of parent-infant inter-brain synchrony, but perhaps even more basic questions about how infant's brains function differently in the presence of an adult/ social partner (e.g., Wass et al., 2020) can help us to develop hypotheses about how two neural systems interact. For example, research with adults has shown that participants spend more time looking at the face of a confederate during a pre-recorded video vs live

(Cañigueral & Hamilton, 2019b). Whilst playing games, adult participants show increased prosocial behaviour under belief of being watched (Cañigueral & Hamilton, 2019b). Another study with 14-month-old infants found that interpersonal movement synchrony increased infants' prosocial behaviour (Cirelli et al., 2014). At the neural level adult's processing of faces has been shown to be different with a live partner vs viewing a picture (Pönkanen et al., 2011a). Further infant's object processing has been shown to be impacted by the presence of a social partner (adult) (Wass et al., 2018). Despite studies like these pointing to clear social influences on behaviour and brain activity very little is understood about the mechanisms that underlie real-time social behaviours (Cañigueral et al., 2022).

The results of the present study show at a very broad level that infant's brains do respond differently to social vs non-social information (as indexed by increased ERP amplitudes to faces vs objects - see figure 1). However, we found no evidence for changes in brain activity around moments when a partner reciprocates mutual gaze (see receiver gaze onsets, chapter 7) which was surprising given the wealth of literature suggesting early development of sensitivity to eye contact (e.g., Farroni et al., 2002). Perhaps, though consideration of the *context* of gaze is key to a better understanding of the basic neural mechanisms that support social behaviours, like gaze. For example, studies have shown that patterns of gaze directed towards a partner differ depending on who is speaking/ listening (Cañigueral et al., 2021) and whether a conversation takes place over a live video call or not (Mansour & Kuhn, 2019).

Perhaps in our study we would have been able to gain a more in depth understanding of the dyadic mechanisms associated with mutual gaze by focussing on the context. To increase the number of trials in our analysis our mutual gaze onsets we included looks that were directed from an object or from a state of inattention. However, it is worth considering that these

likely represent qualitatively and socially distinct experiences, perhaps evoking distinct neural patterns both intra- and interpersonally. This is a possibility that future research should explore in more detail.

An additional way in which future research could consider context in order to further our understanding of the neural dynamics associated with real time social behaviours would be through examining effects of different social partners. To examine whether the (lack of) evidence for inter-brain synchrony in our study depended on partner familiarity we performed additional analyses with infant-adult (unfamiliar to the infant) dyads vs infant-caregiver dyads. We found the same set of results for infant-adult dyads as we did for infant-caregiver dyads, suggesting that the neural mechanisms we were aiming to investigate did not depend on partner familiarity. However, we only considered inter-brain dynamics and did not examine in detail whether behaviourally patterns of gaze directed towards a partner's face differed during infant-adult vs infant-caregiver social interactions. This is a prospect that could create future testable hypotheses on the mechanisms that underpin real time social gaze.

Expanding the variables, we extract from social interactions. Recognising that interacting neural systems are grounded within interacting bodies (Hamilton, 2021) it will be important for future research to take more multivariate approaches. This means increasing the repertoire of behavioural data that researchers collect from their participants. Using high resolution video recordings is a cheap and easy way to do this, as there are now an increasing number of technologies that allow for the tracking of postural movements (hands, face, limbs, head) from pre-recorded videos (e.g., see MediaPipe; Lugaresi et al., 2019). Examining the

temporal dynamics of behavioural coordination across multiple modalities will naturally create testable hypothesis that can be explored within interacting brains. For example, one recent paper (Hale et al., 2020) examined coordination in the head movements of participants during structured conversations, distinguishing between zero-lag concurrent synchrony, fast (e.g., <1s lag) reactive sequential synchrony or slow, sequential synchrony (e.g., over a lag of several seconds). The results show, consistent with the fast reactive model of sequential synchrony that head movements were best modelled by a mechanism with a constant 600ms lag. Similarly using cross-recurrence quantification analysis – a method that is capable of extracting the time lags of peak associations between two time series (similar to a cross correlation) - López Pérez and colleagues (2017), demonstrated patterns of parent-infant behavioural coordination, which varied in the peak time lag of the association dependent upon the behavioural modality and task demands.

Our results show that for eye contact, neural responses in the 'senders' (i.e., the person turning to look at the other person, creating eye contact) brain activity occur rapidly and reactively (< 100ms after the offset of the eye movement). We found no evidence in this work for neural activity in either the sender's or the receiver's brain that preceded moments of mutual gaze which might indicate some degree of anticipation or prediction. Therefore, in our analysis we focused on quantifying inter-brain synchrony around mutual gaze as a concurrent relationship that involved changes in both partner's brain activity at the onset of the eye contact. However, there is ample previous research using screen-based tasks that has shown neural correlates of saccade planning in infancy (Csibra et al., 2000; Richards, 2001). If similar neural correlates can be identified using naturalistic social paradigms it will be interesting for future research to investigate potential (bidirectional) relationships between

this form of action preparation in one individual and the behavioural/ neural responses in a partner following the action.

Bridging the gap between EEG and fNIRS hyperscanning. Largely independently EEG and fNIRS hyperscanning as fields have continued to grow in recent years. However, because of differences in spatial and temporal recording accuracy it can be difficult to interpret whether similar findings across these different neuroimaging methods reflect the same neural mechanisms. For example, Piazza and colleagues (2020) used fNIRS hyperscanning to show during live social interactions eye contact was associated with increased correlation in adult-infant pre-frontal cortex activity. This was compared to moments during which there was no eye contact. Similarly, using fNIRS hyperscanning Hirsch and colleagues (2017) and Noah and colleagues (2020) found live eye contact to be associated with increased inter-brain coherence compared to eye contact with a picture (Hirsch et al., 2017) and eye contact with a pre-recorded video (Noah et al., 2020). However, for studies that contrast eye contact during live interactions with non-live conditions it is quite possible that the mechanisms under investigation are quite different to those in studies that contrast live direct vs live averted eye contact. It is possible that the effects observed in these studies have little to do with eye contact but rather the presence or absence of a live partner (e.g., Hamilton, 2016; Hamilton & Lind, 2016; Cañigueral et al., 2022).

Notwithstanding the differences in the study design, together these findings suggest that there must be neural processes/ structures that are recruited differentially in facilitating direct eye contact. But how can future EEG research investigate the mechanisms by which these patterns of activity become synchronised between individuals? And how can this be informative to researchers investigating similar concepts using fNIRS hyperscanning?

Firstly, is inter-brain synchrony substantiated by two brains mutually adapting to one another? Or is it result of one brain adjusting its activity to be more similar to the other. Using gaze as an example, to explore this researchers will need to carefully consider activity over a range of time lags before and after the onsets of gaze, considering a) if there are patterns of neural activity that precede eye movements in the sender mutual gaze onsets, this might suggest that changes in brain activity associated with mutual gaze start much earlier, during the preparation of the eye movement. b) Under what circumstances and at which temporal lags do receivers anticipate that their partner is going to look at them? Does this always involve patterns of neural activity that bring two individuals closer to alignment, regardless of whether this is desirable or not. These types of questions which require precise temporal information are naturally well suited to EEG research but will provide a foundation of knowledge that will be beneficial for event locked fNIRS study designs also.

Second how do onsets of shared behavioural events like mutual gaze impact the subsequent behaviours of both individuals? And do changes in behaviour after an event result in changes in both partner's neural activity which might affect inter-brain synchrony in the time window following the shared behavioural event? For example, recent research has shown in interacting adults that when one person is looked at they increase the frequency at which they change their gaze (Dobre et al., 2021). We know that eye movements produce substantial increases in EEG power, both as a result of residual artifact and genuine neural activity (e.g., Haresign et al., 2021). EEG power is heavily linked to many measures of inter-brain synchrony (e.g., Muthukumaraswamy et al., 2011). Therefore, it will be important for future research to investigate how changes in behaviour following shared events manifest in changes in intra-individual EEG activity which might also contribute to differences in inter-

brain synchrony. This will be an important point for fNIRS based hyperscanning to consider also as it suggests that inter-brain synchrony might be sensitive to the amount of contingency within and interaction, rather than just its presence or absence.

Lastly, identifying shared loci of inter-brain synchrony between EEG and fNIRS research will allow for more robust validations of the neural systems that being measures through inter-brain synchrony. An important part of this might also be co-registration of EEG and fNRIS signals. If robust correspondences can be obtained between fNIRS and EEG signals (e.g., Pinti et al., 2021) then it should be entirely possible to use EEG to explore fine-grained mechanisms associated with patterns of inter-brain synchrony observed in fNIRS hyperscanning research.

Summary

Overall, the methodological and empirical work presented within this thesis furthers our understanding of the neural dynamics of social interactions in several ways. Firstly, they illustrate that close resemblance and replication of findings from visually evoked ERPs obtained using traditional screen-based experiments is possible using unstructured naturalistic study designs when careful attention (and even using specifically designed solutions) is paid to eye movement artifact within the data, which will by naturally covary with key cognitive/ attentional processes. This is encouraging for the field of cognitive/ social neuroscience broadly given recent drives for more ecologically valid social neuroscience. Second what we

hoped to understand about if/ how inter-brain synchrony arises during social interactions around moments of mutual gaze we were only partially able to achieve.

Chapter 7 details our investigation of inter-brain synchrony around moments of naturally occurring parent-infant mutual gaze. Through analysis of neural responses around moments of mutual gaze in both parents and infants, we aimed to show a chain of evidence that would implicate mutual gaze onsets as a driver of intra-brain phase resetting and subsequently increased inter-brain synchrony. We found no evidence for increased inter-brain synchrony around mutual gaze limiting the extent to which we could evaluate the steps in our chain e.g., whether phase resetting was a key mechanism for establishing inter-brain synchrony. Despite the null findings the results presented in chapter 7 themselves tell us several important things about the neural dynamics of social interactions. Firstly, they show that neural responses to faces contingent on saccades (see sender neural responses in chapter 7) are much greater (in terms of observable changes in amplitude) than neural responses to faces contingent on fixations (see receiver neural responses in chapter 7) which are very subtle/ non-detectable using the experimental set up and design we used. Secondly, they show that if patterns of inter-brain synchrony are a neural index of cognitive processes that support/ facilitate social interaction, for visual attention at least, they are not like other visually evoked neural responses which produce substantial and clearly observable changes in the EEG at the onset and immediately following some event.

The result of our analyses suggest that patterns of inter-brain synchrony cannot be observed in the ways that traditional visually evoked potential can be. Therefore, if inter-brain synchrony is important for social interaction and we want to understand the mechanisms that give rise to it we need think carefully about; a) does inter-brain synchrony involve different

neural structures than those that we know are involved in traditional (screen based) visual/ auditory processing? If so, what's causing these structures to respond differently? B) what do the measures that we are using imply about the properties of underlying signal. For example, if greater phase locking is observed between two people during moments of mutual vs non-mutual gaze, then at some point there would have to have been a change in phase (or frequency) of at least one person's EEG. We need to come up with rational hypotheses for behaviour's that could actually induce these changes and we need methods that are capable of tracking these changes over large periods of continuous data. This can be achieved through consideration of what is already known from screen based social neuroscience (e.g., neural responses to different behaviours), coupled with the use of simulated and empirical data to model potential mechanisms.

Beyond the study of inter-brain synchrony this thesis raises several important areas for future investigation of the neural correlates of real-world gaze/ vision. Are neural responses contingent on saccades different to those contingent on fixation? If so, why? How much of what we process visually about our environment happens immediately following saccades? When and where does processing of a partner's gaze shift take place? Does neural (motor) activity preceding saccades carry information about the saccade, e.g., where and when it will occur? And how long will the look last for?

To summarise the result of the present thesis, challenge the conclusions drawn from both traditional screen-based investigation of eye processing in infancy (e.g., Farroni et al, 2002) . First, when we repeat the Farroni et al analysis, we find no evidence to support the hypothesis that, during real-world naturalistic social exchanges, infant's brains respond differently when someone looks directly at them. This is despite the fact our data are clean, as shown by the

clear face/object ERPs that we include in the SM, and the other data quality metrics that we also report in the SM. Our sample size was also larger than in the original study (N=55 vs N=16 in the original study).

Second, when we examine inter-brain entrainment during direct vs averted gaze, we also find no evidence to replicate the previous finding, of Leong and colleagues (2017) that Granger-predictive associations in brain activity are stronger during direct gaze. This is, again, despite that our sample size was larger than in the original study (N=55 vs N=17 per group). Instead, our findings suggest a new conclusion: that the effects of mutual gaze are strongest at the intra-brain level, in the 'sender' but not the 'receiver' of the mutual gaze.

The findings presented in this thesis will be crucial in adding empirical evidence to recent theoretical papers that 'push back' against the rapid growth of interest in dyadic neuroimaging recording techniques that is currently taking place around the world. Given that they go to the very heart of our current prevailing theories of how our brains learn to process social information.

Appendices

Appendix A – Ethics application form

Ethics approval for the project was submitted and obtained by Megan Whitehorn on behalf of the three PhD students working on the project; Myself, Megan Whitehorn and Emily Phillips. Below is the approved ethics application letter.

Ethics ETH1819-0141: Ms Megan Whitehorn (High risk)

Date	10 Apr 2019
Researcher	Ms Megan Whitehorn
Student ID	1920537
Project	The effects of Attention Training on Neural and Behavioural Responsivity in Infants During Parental Dyadic Play
School	Psychology

Ethics application

Checklist for research projects conducted during a pandemic

Project details

1.1 Is your research project taking place during a pandemic?

1. Project details

1.1. Title of proposed research or consultancy project

New insights into how the infant brain subserves dynamic social interactions

1.2. UEL Researchers or Consultancy lead'

Ms Megan Whitehorn

Ms Emily Phillips

Mr Ira Marriott Haresign

Prof Samuel Wass

1.3. Start date of project for which ethical approval is being sought

24 Jun 2019

1.4. Anticipated end date of project for which ethical approval is being sought 21 Jan 2022

1.5. If this project is part of wider research or consultancy work, please provide the UREC, EISC, URES, RRDE, SREC, CREB or NHS research ethics approval number

n/a

1.6. If this project is part of a wider research study or consultancy work please state the start and end dates n/a

1.7. Specify where the research or consultancy project will take place

UEL, Stratford Campus

2. Aims and methodology

2.1. Aims and objectives of the project

Most early cognitive learning takes place in the presence of an adult social partner. Behavioural and psychophysiological research has suggested that as infants and parents communicate they adapt to each-other on a moment-by-moment basis- through, for example, cycles of vocalizing and pausing, of looking towards and away from each other, and of matching each-others' positive, negative and neutral affect (Feldman, 2007). Less well

understood is whether similar mechanisms of mutual attunement also operate at the neural level, and if/how shared patterns of oscillatory activity between individuals sub-serve interpersonal influences on attention.

Using dual-EEG recordings, recent work has found that when interacting with an adult partner, increased neural synchronisation is observed during moments of direct gaze (Leong et al., 2017). A further study conducted with mothers and infants has also shown that when parents engage in joint play with their infant, parents' neural activity tracks and responds to infant attention, with greater responsivity associated with increased infant attentiveness. The primary aim of the project is to further our understanding of interpersonal mechanisms that give rise to parent-infant neural "phase locking" (a phenomenon whereby two oscillatory signals become temporally aligned). Our secondary aim is to investigate the effects of child-focussed vs. parentfocussed intervention on parent-child behavioural synchrony, and how these effects are substantiated at the neural and physiological level. The child-focussed intervention group will undergo a computerised attention training intervention, previously shown to lead to increased voluntary attentional control and responsiveness to social cues (Wass et al.; Wass & Forssman, 2017). The parent-focussed intervention group will receive a parent-child book-sharing training programme. This intervention has been shown to increase parent responsivity to infant cues, leading to gains in child attentional control, as well as receptive and expressive language abilities (Valley et al., 2016). By comparing parent-child interactions in the two groups before and after intervention, we aim to identify casual mechanisms to behavioural, physiological (ECG) and neural (EEG) synchrony.

2.2. Methodology, data analysis and recruitment for the project

To assess neural phase locking, we shall record simultaneous Electrocardiography (ECG) and dual Electroencephalography (EEG) from N=210 typically developing 10-12-month-old infants and adults before and after intervention.

Participation in the study will involve attending 8 sessions, which will be scheduled weekly. The first 2 visits will be pre-intervention sessions; these will be identical for all participating infants. Then, the subsequent 5 visits will be intervention visits. The exact schedule for these will vary depending on which group the participant is in (see further details below). The last visit will be a post-intervention session. These will, again, be identical for all participants. Allocation of participants to intervention group will be fully randomised and performed prior to the participants' first testing session.

Infants will be recruited from the participant database at UEL, as well as various local 'baby-groups', following the same procedures as used in multiple previous studies. Participants will also be recruiting from our pre-existing participant database. All participants will be aged between 270 and 330 days at the date of the first visit. Exclusion criteria will be: complex medical conditions, skin allergies, heart conditions, parents below 18 years of age, and parents receiving care from a mental health organisation or professional.

Pre-intervention assessment (visits 1 and 2). The infant's parent will be with them at all times. During these pre-intervention visits, both parents and their parents will have their electrical brain activity measured using Electroencephalography (EEG) and Electrocardiography (ECG). EEG will be measured using a Biosemi 32-channel system optimised for dual EEG recording. This is a gel-based system that is already in use in

multiple sites across the world for infant testing. ECG will be measured using stick-on electrodes placed in a modified lead II position.

Participants will complete three conditions, spread across two visits:

In Condition 1 (Video), infants will view a pre-recorded adult experimenter continually reciting nursery rhymes. Adult's pre-recorded EEG will be compared with the infant's live EEG during viewing.

Behavioural data will be videoed and coded post-hoc.

In Condition 2 (Live), we shall record dual EEG and ECG from an infant and an unfamiliar adult (researcher) simultaneously, while they engage in table-top play. Behavioural data will be videoed and coded post hoc.

In Condition 3, the same procedure will be repeated while the infant interacts with their parent, we shall record dual EEG and ECG from an infant and and parent simultaneously, while they engage in table-top play. Behavioural data will be videoed and coded post hoc.

Participants will also be administered a battery of cartoon-based cognitive and language outcome measures using eyetracking. The eyetracker used will be a Tobii TX300, that has been used in numerous studies around the world. Parent questionnaires will also be administered, namely the: MacArthur Communicative Development Inventory; the Infant Behaviour Questionnaire (short version); the GAD-7 and PHQ-9 (clinical assessments of parental anxiety and depression symptoms); the Penn State Worry Questionnaire, and a demographics questionnaire. All questionnaires proposed for inclusion are given in the appendix. Sessions will last a total of 2 hours. This is a standard length of time for lab visits with infants of this age; although assessments will last around 30-60 minutes, we like to allow enough time for the infant to settle in, and have breaks (for naps, snacks etc).

Intervention 1- Child focused attention training intervention.

This intervention will be delivered once a week over a five-week period, with each session lasting 30 (+/- 10) minutes. Each session will involve the child being positioned on their parent's lap in front of an eye-tracker, while training stimuli are presented on a computer monitor. Training stimuli consist of 6 tasks presented consecutively in the same order for each session; different events take place in the stimuli contingent on where the infant is looking. The intervention has already been published (Wass et al., 2011, 2018) and been used in several studies without any problems or unintended negative outcomes being reported.

Intervention 2 – Parent focussed book sharing intervention.

This intervention trains parents in 'dialogic' book-sharing techniques; active and evocative behaviours that engage children in reciprocal communication during book-sharing. The intervention will be delivered once a week over five weeks by one of the PhD students on a one-to-one basis, with each session lasting 60 +/- 10 minutes. Each week will involve a specific theme, e.g. 'elaborating and linking', as well as a book-of-the-week for parents to take home and practice the book-sharing techniques with their children. The intervention has already been published (Vally et al., 2015) and been used in several studies without any problems or unintended negative outcomes being reported.

Intervention 3 – Control

Infant in the control group will make similar lab visits to the intervention groups; they will watch some computerised animations, and the researcher will discuss ways to incorporate foods into the child's diet that may be particularly helpful to supporting their early development.

Post Intervention

Following training, all participants will attend two post-assessments. These assessments will follow exactly the same format as the pre-assessment sessions. Participants will receive the same battery of pre-assessment outcome measures.

Through analysis we aim to explore interpersonal neural and physiological correlates of social interaction. We will compare parent-child behavioural, physiological and neural synchrony before and after intervention, in order to investigate the effects of receiving each intervention, and in comparison to no intervention at all.

2.3. Is the data accessed, collected or generated of a sensitive nature?

Yes

2.3.1. If yes, please provide details. Please ensure that all data of a sensitive nature is handled carefully and stored appropriately.

We will be collecting heart rate and EEG data. These will be stored securely on a passwordprotected drive kept under lock and key in the university and accessible only by project researchers.

3. About your project

3.1. Is the research/consultancy project funded?

Yes

3.2. Does the project involve external collaborators?

No

3.3. Does the project involve human participants?

Yes

3.4. Does the project involve non-human animals? No

3.4.1. If yes, where is the research project taking place?

3.5. Does your project involve access to, or use of, material (including internet use) covered by the Terrorism Act (2006) and / or Counter-Terrorism and Border Security Act (2019) or which could be classified as security sensitive?

No

3.6. Does the project involve secondary research, secondary data or analysing an existing data set? No

3.7. Does the project raise ethical issues that may impact on the natural environment over and above that of normal daily activity?

No

3.8 Does the research/consultancy project involve data collected online via social media, advertising the project online or via social media or include a questionnaire/survey?

No

If yes, please provide details.

3.9. Will the research/consultancy project take place overseas?

No

3.10. Will the researcher or research team be responsible for the security of all data collected in connection with the research/consultancy project? Yes

3.11. Does your research/consultancy project require third-party permission?

No

If yes, please provide details.

3.12. Does your research/consultancy project involve any circumstances where the professional judgement of you and/or the team is likely to be influenced by personal, institutional, financial or commercial interests?

If yes, please provide details.

3.13. Does the project involve consultancy or contract research?

If yes, please provide details.

4. Funding

4.1. Funder(s)

Leverhulme Trust

4.2. Grant type

Research Council

If you selected other, please provide further details.

4.3. Value of grant £ 327093

4.4. Please upload a letter advising of the award of the grant.

6. Recruitment

6.1. Are the research participants able to give informed consent (in written or verbal form)?

No

6.1.1. If no, is this because they are perceived to lack mental capacity or because they are vulnerable? Vulnerable

6.1.2. If the participants are perceived to lack mental capacity, please provide the reason(s).

6.1.3. Further details

6.1.4. If the participants are perceived to be vulnerable, please provide details of the vulnerability.

Research will involve children aged 10-12 months at time of training and testing. In line with previous studies of this type undertaken by the university, parents will be asked to consent to each test proposed in this study. Written information about each test will be emailed to parents at least a week before their visit to the lab, and they will be encouraged to ask for any

clarification they feel they need. All participants are invited to ask questions and voice concerns about our consent and information documents, so that we can respond to or expand on any part of the process that is not clear.

6.1.5. Does the research/consultancy project involve children or young people under the age of 16? Yes

6.1.6 If yes, are the children or young people able to give informed assent?

No

6.1.7. If no, is this because they are perceived to lack mental capacity or because they are vulnerable? Vulnerable

6.1.8. If the participants are perceived to lack mental capacity, please provide the reason(s).

6.1.9. Further details

6.1.10. If the participants are perceived to be vulnerable, please provide details of the vulnerability.

The infant participants will be aged between 10 and 12 months at the time of training and testing. At these ages, the babies will be too young to give informed consent before starting, therefore parents will consent on behalf of their infants/children, as standard in this age cohort. In addition, they will be constantly monitored for signs of distress, and any procedure that is deemed upsetting to either parent, researcher or the baby will be halted. All participants will be informed that they have the right to halt proceedings at any stage. Babies and children will be constantly monitored, and any test will be discontinued should they become distressed or show signs of discomfort. This includes both training and testing sessions.

6.2. How will participants be recruited?

Infants will be recruited from the participant database at UEL, as well as from various recruitment drives at local baby-groups and Children's Centres. Fully informed consent will be obtained before testing commences.

All of the parents of babies in this study will have received an information sheet and been given an explanation of the aims of this study before providing their contact details. Different information sheets will be given depending on which group the participants have been pre-allocated to. When they are initially contacted, they will be given more detailed information about what participation in the study involves via an information sheet, which will be emailed to them. At each interaction, an opportunity will be given to ask any questions or gain clarification. Before any data is collected, the parent/carer will be asked to sign a consent form. All participants are invited to ask questions and voice concerns about our consent and information documents, so that we can respond to or expand on any part of the process that is not clear.

6.3. Please upload recruitment documents.

6.4. How many participants are being sought for the project?

210

6.5. How long will participants be required for the project?

The intervention blocks will take place over 9 weeks, with testing sessions once each week. The pre- and post- assessments will each last for two hours. Intervention sessions will last 30 +/- 10 minutes for the attention training and 60 +/- 10 minutes for the book-sharing.

6.6. Will the participants be remunerated for their contribution?

Yes

6.6.1. If yes, please specify monetary value of cash or giftcard / vouchers.

£50 in shopping vouchers for those who receive the interventions and £30 in vouchers for those who do not.

7. DBS

7.1. Do you require Disclosure Barring Service clearance (DBS) to conduct the research/consultancy project?? Yes

7.2. Is your DBS clearance valid for the duration of the research/consultancy project?

Yes

7.2.1. If you have current DBS clearance, please provide your DBS certificate number.

001659510034

8. Medical

8.1. Is your project a clinical trial and / or involves the administration of drugs, substances or agents, placebos or medical devices? No

8.1.1. If you answered yes, please explain why you have chosen to use this application form instead of the NHS/HRA ethics application form. If you have selected yes, your project requires approval by the NHS/HRA, as it is falls under the classification of Medicines for Human Use (Clinical Trials) Regulations (2004) or Medical Devices Regulations (2002) or any subsequent amendments to the regulations.

8.2. Does your project involve the collecting, testing or storing of human tissue / DNA including organs, plasma, serum, saliva, urine, hair, nails or any other associated material?

No

8.2.1. If you answered yes, please explain why you have chosen to use this application form instead of the NHS/HRA ethics application form. If you have selected yes your project requires approval by the NHS/HRA, as it is falls under the classification of the Human Tissue Act (2004).

9. Risk

9.1. Does the project have the potential to cause physical or psychological harm or offence to participants and / or researchers?

No

9.1.1. If yes, please provide details of the risk or harm explaining how this will be minimised.

9.1.2. Please complete and upload a research risk assessment form

9.2. Does the project involve potential hazards and/or emotional distress?

Yes

9.2.1. If yes, provide an outline of support, feedback or debriefing protocol.

Infants sometime express mild distress on the application of the EEG equipment. However, we have received extensive training in how to minimise this. We do not obtain any usable EEG recordings if infants are at all distressed, so we take every possible step to minimise this when it occurs. The process of recording EEG data from infants is in place in numerous other research labs across the world, and the process we follow will be identical to those used in other labs.

9.3. Provide an outline of any measures you have in place in the event of an adverse event or reaction or unexpected outcome, the potential impact on the project and, if applicable, the participants.

We do not expect adverse events arising from the study procedures. Under circumstances of an unexpected adverse event, the participants' health and safety will be the highest priority. If health and safety is in any way compromised, the participant will be withdrawn from the study, with clear reasoning given.

10. Anonymisation

10.1. Will the participants be anonymised at source?

No

10.1.1. If yes, please provide details of how the data will be anonymised.

10.2. Are participants' responses anonymised or are an anonymised sample?

No

10.2.1. If yes, please provide details of how the data will be anonymised.

10.3. Are the samples and data de-identified? Yes

10.3.1. If yes, please provide details of how the data will be anonymised/pseudonymised.

Direct and indirect identifiers will be removed from data and participants will be assigned a participant code. This will be entered along with the date and time of testing on the files containing raw data, and will be used to record all other data collected during pre-post assessments. Participant names and ID numbers will be stored in a separate password protected database. This is so that participants can withdraw their data up to the point at which it is included in the final analysis. Consent forms will be kept securely and separately from the raw data in a locked cabinet. Only members of the research team will have access to both the raw data and consent forms. Data will be retained in a secure place at the end of the project as, should funding allow, we might wish to follow up the sample over a longer time period. Video recording of the behavioural paradigms will only be viewed by members of the research team.

10.4. Please provide details of data transcription.

All data will be analysed using MATLAB software. Behavioural paradigms will be coded according to their manuals. External research assistants may be recruited in order to code videos; assistants will be fully briefed on ethics related to the study and will not be allowed access to consent forms, code keys, or any other data that would allow them to identify the participant. Research assistants will only be permitted to code data on university computers (that are disconnected from the internet) and will not be allowed to copy or move video files

from the university. Research assistants will be asked to complete a confidentiality agreement before undertaking any work on the project.

10.4.1 Will the data be transcribed by person(s) outside of the project team?

If yes, please upload a blank copy of the confidentiality agreement.

10.5. If applicable, will all members of the project team know how the code links the data to the individual participant? Yes

10.5.1. If no, in the event of a researcher's absence please specify the process should access to the research data be required.

10.6. Will participants be anonymised/pseudonymised in publications that arise from the research/consultancy project? Yes

10.6.1. If no, please provide details.

10.7. Will participants have the option of being identified in the study and dissemination of research findings and / or publication?

No

10.7.1. If yes, please provide details.

16. Data security

16.1. Will the researcher or the project team be responsible for the security of all data collected in connection with the proposed research/ consultancy project? Yes

16.1.1. If no, please provide details.

16.2. Will the research/consultancy data be stored safely on a password protected computer? Yes

16.2.1. If no, please provide details.

16.3. Will the research/consultancy project data be stored on a UEL data managed device? Yes

16.3.1. If no, please specify where the electronic data will be stored and how the data will be kept secure.

16.4. Will you keep research/consultancy project data, codes and identifying information in a separate location? Yes

16.4.1. If yes, please explain how you will store the research data.

Consent forms will be kept in hard copies in a locked cupboard on university premises. Only the research team will have access to this cupboard. The cupboard will be locked when not in use. Data and audio/video recordings will be stored separately from the consent forms. Hard copies of personal data will be stored in locked cupboards in the lab. Electronic versions of personal data will be stored on a secure computer network (to which only the research team have access) on password protected computers at the University of East London. Participant names and IDs will be stored on a separate password-protected database.

16.5. Will the raw data be shared with individuals outside of the project team?

No

16.5.1. If yes, please specify the names, positions and their relationship to the research/ consultancy project

---

Name

Position

Relationship to research

16.6. Will participants be audio and/or video recorded?

Yes

16.6.1. If yes, please explain how you will transfer, store and, where relevant, dispose of audio and/or video recordings.

Digital audio-video recordings will be transferred onto the UEL secure computer network that only the research team will have access. Video files may need to be stored externally; in this case, recordings will be stored and transferred on a password-encrypted hard drive with access limited to members of the research team.

16.7. If audio and/or video recordings will be retained, please provide details and state how long the recordings will be kept.

Recording will be kept for up to a year after the data has been analysed unless further funding can be obtained and the data is used again in further analyses, for which ethics approval of analyses of secondary data will be sought.

16.8. Will you retain hard copies of the data? Yes

16.8.1. If yes, please provide details of how the data will be transported safely and, where relevant, undergo secure disposal.

Data will be collected and stored in the BabyLab at UEL. Hard copies of any data will be kept in locked cabinets in the BabyLab offices. Only members of the research team will have keys to those cabinets. The data will not leave the premises. These will be shredded by a team member when appropriate.

16.9. Will the research/consultancy project data be encrypted and transferred inside of the

UK?

Yes

16.10. How long will the research data that details personal identifiers be stored?

Any personal identifiers will only be retained until the research has been published and it is no longer possible for participants to request to withdraw their data

16.11. Please upload a copy of your Data Management Plan.

18. Dissemination

18.1. Will the results be disseminated? Yes

18.1.1. If yes, how will the results of the research/consultancy project be reported and disseminated?

Dissertation / Thesis

Peer reviewed journal

Internal report

Conference presentation

18.1.2. If you selected other, please provide further details.

18.1.3. If the results of the research/consultancy /project will not be reported and disseminated, please provide a reason

20. Attachments

You can generate a Participant Information Sheet and Consent Form using the answers provided in your ethics application form. The Word files generated can be edited. You should upload the final version(s) before submitting your application form.

20.1. Upload any additional files to support your application which have not already been uploaded within your application.

Appendix B – Covid-19 Risk Assessment Form

Covid-19 Risk Assessment Form

Use this form for assessing risk specifically related to our work around the management of the Covid-19 outbreak all forms should be stored on Silver Command Teams Site

Risk Owner (who's responsible for managing the risk)	**Dr Sam Wass**	**Date of Assessment**	**12/01/2021**
		Date of Review	
Activity title	**Use of research suites in Arthur Edwards building for testing**	**Location of activity:**	**Arthur Edwards building, Stratford Campus**
Signed off by Manager (Print Name)		**Date and time (if applicable)**	**Various – always during office hours**
Financial Cost to UEL	**See UEL Resources Required section.**	**UEL Resources Required**	**PPE (surgical masks, gloves), disinfectant. If these are not available centrally from the university they can be purchased by us from grant income.**

Please describe the activity/event in as much detail as possible. Please include information about the nature of activity, the resources required, the estimated number of participants and the start and end date.

All of the testing sessions involve individual parents bringing their children to attend sessions in the lab that are run by researchers working either individually or in pairs. No testing session will ever involve more than three adults and one child.

All of our testing involves parents and infants, who generally fall within the low-risk categories for COVID-19 – although see below for the steps we will take to ensure this.

To allow us to reopen testing in a way that minimises the risk to participants and to researchers, and considering that the new variant of COVID 19 is more transmissible, we propose the following steps:

Prior to the session:

- only researchers who are in low-risk groups (i.e. not over 70, no underlying conditions, and not living with anyone who falls into these categories) will be asked to conduct testing in the BabyDevLab suite. Named researchers are: Louise Goupil, Katherine Denman, Zeynep Suata, Emily Phillips, Megan Whitehorn, Ira Marriott Haresign. Testing is run during normal office hours and occasional Saturday.
- on contacting a potential participant we will explicitly ask whether there is any reason why they do not consider themselves low-risk (using the same criteria as above), and we will only schedule visits with low-risk participants. Anyone who is 'clinically extremely vulnerable' will not be allowed come to campus.
- all potential participants and researchers will be asked to self-report as to whether there is a chance that they have been exposed to COVID-19 within the past 14 days (following Public Health England guidelines) and will be asked not to attend testing sessions if they have.
- researchers will take a lateral flow COVID test twice a week where possible and appropriate.
- participants will be strongly encouraged to take a lateral flow test within 3 days prior to their first visit, and if possible send a screen shot of their results.
- participants will be told they need to have an active NHS 'track and trace' or 'COVID-19' app if possible, before visiting the lab.
- on arrival, researchers and participants will be asked to scan the displayed 'track and trace' posters at the UEL main reception. Participants who do not have the app will be asked to write down their details in a dated log. Pens will be disinfected after use.
- researchers and participants will be asked to take their temperature using one of the machines in the main UEL reception, when they arrive on campus.
- researches will log that all mitigation measures have been followed at each visit.
- we will ensure that only one participant attends a testing session in any research room at one time, and that there is at least an hour between visits in order to allow researchers to fully disinfect between visits any surfaces that were touched.
- researchers and participants will only be allowed to travel to the labs by walking, bike or car. If this is not available we will arrange and pay for taxis in and out. (These costs will be covered from grant income.)
- it will be emphasized to participants that travel to the lab should be direct to fall in line with recent Government guidance to 'stay at home'
- researchers will be asked only to travel in and out of the lab for their scheduled testing visits. At all other times, they will continue to work from home.

- for participants who drive in to the lab, we will arrange with Security so that their parking voucher can be provided electronically
- we will ensure that all of the paperwork that participants normally fill in at the start of their visit (i.e. consent forms and questionnaires) is done online, via Qualtrics, in order to keep the visit as short as possible and avoid having to touch pen and paper.
- all briefing of participants, that normally we do when participants first arrive, will instead be done beforehand, by phone.
- rather than providing food and drink and refreshments for the parents, as we normally do, we shall instead ask the parents to bring their own food and drink. We will offer a small financial reimbursement to parents to compensate for this, which will be paid electronically.
- participants who develop symptoms and get tested, or who test positive within two days after visiting the lab, will be asked to notify us, so that we may notify those people who were there at the same time.
- the wider COVID-19 situation will be kept under weekly review and risk mitigations measures revisited as and when necessary.

During the session:

- researchers will thoroughly disinfect all surfaces that will be touched prior to the participants' arrival in the lab. We will ensure that adequate training is provided so that researchers do this effectively.
- participants will be asked to wash their hands thoroughly on arriving at the lab (in the drugs lab), and then provided by us with a surgical mask and gloves that they will wear through their visit (with a few exceptions, as described below).
- researchers will wear surgical mask and gloves throughout the visit (with no exceptions). These will be provided by the university.
- participants will be met by the researcher at reception and will be accompanied by our researchers to the research suites, ensuring that there is no possibility of exposure on the way to the lab.
- most of the equipment we use can be applied by the participants themselves, under the remote direction of the researchers. This means that it will be possible to maintain 2 metres distance between researchers and participants throughout the visit. The only exception to this is the EEG recording nets, that have to applied by the researchers. However, we have discussed this carefully with the testing staff and this can be done always with the researchers behind the participants, applying the cap directly to their heads, without any face-to-face contact.
- Researchers will thoroughly disinfect all surfaces touched during each visit, immediately after the participants leave. At least one hour's gap will be scheduled between visits in order to allow time for this.

All staff

- All staff will work from home wherever possible. When research or other staff come into the office for testing purposes they will follow the University's 'Daily 123' protocol:
  1. TEST - Before coming onto campus each day staff will complete UEL's online health status update to determine

  whether or not you may attend campus that day; 2. DISTANCE – staff will check-in when they arrive and leave campus

  using the free Safezone app. 3. TRACK – staff will download and use the NHS coronavirus contact-tracing.

- In line with the Daily 123 protocol, staff will download and use the Safezone app.

References

Attaheri, A., Choisdealbha, Á. N., Di Liberto, G. M., Rocha, S., Brusini, P., Mead, N., ... & Goswami, U. (2022). Delta-and theta-band cortical tracking and phase-amplitude coupling to sung speech by infants. *NeuroImage*, *247*, 118698.

Baccalá, L. A., & Sameshima, K. (2001). Partial directed coherence: a new concept in neural structure determination. *Biological cybernetics*, *84*(6), 463-474.

Balestrieri, E., Ronconi, L., & Melcher, D. (2022). Shared resources between visual attention and visual working memory are allocated through rhythmic sampling. *European Journal of Neuroscience*, *55*(11-12), 3040-3053.

Barnett, L., Barrett, A. B., & Seth, A. K. (2009). Granger causality and transfer entropy are equivalent for Gaussian variables. *Physical review letters*, *103*(23), 238701.

Barnett, L., Barrett, A. B., & Seth, A. K. (2018). Misunderstandings regarding the application of Granger causality in neuroscience. *Proceedings of the National Academy of Sciences*, 201714497

Barnett, L., Barrett, A. B., & Seth, A. K. (2018). Solved problems for Granger causality in neuroscience: A response to Stokes and Purdon. *NeuroImage*, *178*, 744-748.

Barnett, L., & Bossomaier, T. (2012). Transfer entropy as a log-likelihood ratio. *Physical review letters*, *109*(13), 138105.

Barnett, L., & Seth, A. K. (2014). The MVGC multivariate Granger causality toolbox: a new approach to Granger-causal inference. *Journal of neuroscience methods*, *223*, 50-68.

Barrett, A. B., Murphy, M., Bruno, M. A., Noirhomme, Q., Boly, M., Laureys, S., & Seth, A. K. (2012). Granger causality analysis of steady-state electroencephalographic signals during propofol-induced anaesthesia. *PloS one*, *7*(1), e29072.

Barry-Anwar, R. A., Burris, J. L., Estes, K. G., & Rivera, S. M. (2017). Caregivers and strangers: The influence of familiarity on gaze following and learning. *Infant Behavior and Development*, *46*, 46-58.

Bart Krekelberg (2022). bayesFactor (https://github.com/klabhub/bayesFactor), GitHub. Retrieved February 28, 2022.

Begus, K., Gliga, T., & Southgate, V. (2014). Infants learn what they want to learn: Responding to infant pointing leads to superior learning. *PloS one*, *9*(10), e108817.

Begus, K., Gliga, T., & Southgate, V. (2016). Infants' preferences for native speakers are associated with an expectation of information. *Proceedings of the National Academy of Sciences*, *113*(44), 12397-12402.

Berens, P. (2009). CircStat: a MATLAB toolbox for circular statistics. *J Stat Softw*, *31*(10), 1-21.

Biau, E., Torralba, M., Fuentemilla, L., de Diego Balaguer, R., & Soto-Faraco, S. (2015). Speaker's hand gestures modulate speech perception through phase resetting of ongoing neural oscillations. *Cortex*, *68*, 76-85.

Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K. M., & Robbins, K. A. (2015). The PREP pipeline: standardized pre-processing for large-scale EEG analysis. *Frontiers in neuroinformatics*, *9*, 16.

Billings, R. J. (1989). The origin of the occipital lambda wave in man. *Electroencephalography and clinical neurophysiology*, *72*(2), 95-113.

Bloom, K. (1988). Quality of adult vocalizations affects the quality of infant vocalizations. *Journal of child language*, *15*(3), 469-480.

Bögels, S. (2020). Neural correlates of turn-taking in the wild: Response planning starts early in free interviews. *Cognition*, *203*, 104347.

Breton, A., Jerbi, K., Henaff, M. A., Cheylus, A., Baudouin, J. Y., Schmitz, C., ... & Van der Henst, J. B. (2014). Face the hierarchy: ERP and oscillatory brain responses in social rank processing. *PloS one*, *9*(3), e91451.

Burgess, A. P. (2012). Towards a unified understanding of event-related changes in the EEG: the firefly model of synchronization through cross-frequency phase modulation. *PloS one*, *7*(9), e45630.

Burgess, A. P. (2013). On the interpretation of synchronization in EEG hyperscanning studies: a cautionary note. *Frontiers in human neuroscience*, *7*, 881.

Busch, N. A., Dubois, J., & VanRullen, R. (2009). The phase of ongoing EEG oscillations predicts visual perception. *Journal of neuroscience*, *29*(24), 7869-7876.

Bushnell, I. W. R., Sai, F., & Mullin, J. T. (1989). Neonatal recognition of the mother's face. *British journal of developmental psychology*, *7*(1), 3-15.

Buzsáki, G., Anastassiou, C. A., & Koch, C. (2012). The origin of extracellular fields and currents—EEG, ECoG, LFP and spikes. *Nature reviews neuroscience*, *13*(6), 407-420.

Cañigueral, R., & Hamilton, A. F. D. C. (2019). Being watched: Effects of an audience on eye gaze and prosocial behaviour. *Acta psychologica*, *195*, 50-63.

Cañigueral, R., Krishnan-Barman, S., & Hamilton, A. F. D. C. (2022). Social signalling as a framework for second-person neuroscience. *Psychonomic Bulletin & Review*, 1-13.

Cañigueral, R., Ward, J. A., & Hamilton, A. F. D. C. (2021). Effects of being watched on eye gaze and facial displays of typical and autistic individuals during conversation. *Autism*, *25*(1), 210-226.

Canolty, R. T., & Knight, R. T. (2010). The functional role of cross-frequency coupling. *Trends in cognitive sciences*, *14*(11), 506-515

Chaumon, M., Bishop, D. V., & Busch, N. A. (2015). A practical guide to the selection of independent components of the electroencephalogram for artifact correction. *Journal of neuroscience methods*, *250*, 47-63.

Cirelli, L. K., Einarson, K. M., & Trainor, L. J. (2014). Interpersonal synchrony increases prosocial behavior in infants. *Developmental science*, *17*(6), 1003-1011.

Cohen, M. X. (2014). *Analyzing neural time series data: theory and practice*. MIT press.

Conte, S., Richards, J. E., Guy, M. W., Xie, W., & Roberts, J. E. (2020). Face-sensitive brain responses in the first year of life. *NeuroImage*, *211*, 116602.

Conty, L., N'Diaye, K., Tijus, C., & George, N. (2007). When eye creates the contact! ERP evidence for early dissociation between direct and averted gaze motion processing. *Neuropsychologia*, *45*(13), 3024-3037.

Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & Language*, *25*(2), 141-168.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in cognitive sciences*, *13*(4), 148-153.

Csibra, G., Tucker, L. A., Volein, Á., & Johnson, M. H. (2000). Cortical development and saccade planning: the ontogeny of the spike potential. *Neuroreport*, *11*(5), 1069-1073.

Cuevas, K., Cannon, E. N., Yoo, K., & Fox, N. A. (2014). The infant EEG mu rhythm: methodological considerations and best practices. *Developmental Review*, *34*(1), 26-43.

de Graaf, T. A., & Duecker, F. (2022). No effects of rhythmic visual stimulation on target discrimination: An online alpha entrainment experiment. *European Journal of Neuroscience*, *55*(11-12), 3340-3351.

De Haan, M., & Nelson, C. A. (1997). Recognition of the mother's face by six-month-old infants: A neurobehavioral study. *Child development*, *68*(2), 187-210.

De Haan, M., & Nelson, C. A. (1999). Brain activity differentiates face and object processing in 6-month-old infants. *Developmental psychology*, *35*(4), 1113.

Debnath, R., Buzzell, G. A., Morales, S., Bowers, M. E., Leach, S. C., & Fox, N. A. (2020). The Maryland analysis of developmental EEG (MADE) pipeline. *Psychophysiology*, *57*(6), e13580.

Delorme, A. (2022). EEG is better left alone. *bioRxiv*.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods*, *134*(1), 9-21.

Dikker, S., Silbert, L. J., Hasson, U., & Zevin, J. D. (2014). On the same wavelength: predictable language enhances speaker–listener brain-to-brain synchrony in posterior superior temporal gyrus. *Journal of Neuroscience*, *34*(18), 6267-6272.

Dimigen, O. (2020). Optimizing the ICA-based removal of ocular EEG artifacts from free viewing experiments. *NeuroImage*, *207*, 116117.

Dimigen, O. (2021). Fixation-related potentials in total darkness. *Journal of Vision*, *21*(9), 2931-2931.

Dimigen, O., Sommer, W., Hohlfeld, A., Jacobs, A. M., & Kliegl, R. (2011). Coregistration of eye movements and EEG in natural reading: analyses and review. *Journal of experimental psychology: General*, *140*(4), 552.

Dimigen, O., Valsecchi, M., Sommer, W., & Kliegl, R. (2009). Human microsaccade-related visual brain responses. *Journal of Neuroscience*, *29*(39), 12321-12331.

Ding, N., & Simon, J. Z. (2014). Cortical entrainment to continuous speech: functional roles and interpretations. *Frontiers in human neuroscience*, *8*, 311.

Dobre, G. C., Gillies, M., Falk, P., Ward, J. A., Hamilton, A. F. D. C., & Pan, X. (2021, October). Direct gaze triggers higher frequency of gaze change: An automatic analysis of dyads in unstructured conversation. In *Proceedings of the 2021 International Conference on Multimodal Interaction* (pp. 735-739).

Doelling, K. B., Arnal, L. H., Ghitza, O., & Poeppel, D. (2014). Acoustic landmarks drive delta–theta oscillations to enable speech comprehension by facilitating perceptual parsing. *Neuroimage*, *85*, 761-768.

Dubey, I., Georgescu, A. L., Hommelsen, M., Vogeley, K., Ropar, D., & Hamilton, A. F. D. C. (2020). Distinct neural correlates of social and object reward seeking motivation. *European Journal of Neuroscience*, *52*(9), 4214-4229.

Dumas, G. (2011). Towards a two-body neuroscience. *Communicative & integrative biology*, *4*(3), 349-352.

Dumas, G., Nadel, J., Soussignan, R., Martinerie, J., & Garnero, L. (2010). Inter-Brain Synchronization during Social Interaction. *PLoS ONE, 5*(8), e12166.

Elsabbagh, M., Volein, A., Csibra, G., Holmboe, K., Garwood, H., Tucker, L., ... & Johnson, M. H. (2009). Neural correlates of eye gaze processing in the infant broader autism phenotype. *Biological psychiatry*, *65*(1), 31-38.

Faes, L., Erla, S., Porta, A., & Nollo, G. (2013). A framework for assessing frequency domain causality in physiological time series with instantaneous effects. *Philosophical*

Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, *371*(1997), 20110618.

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National academy of sciences*, *99*(14), 9602-9605.

Farroni, T., Massaccesi, S., Pividori, D., & Johnson, M. H. (2004). Gaze following in newborns. *Infancy*, *5*(1), 39-60.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, *39*(2), 175-191.

Feldman, R. (2007). Parent–infant synchrony and the construction of shared timing; physiological precursors, developmental outcomes, and risk conditions. *Journal of Child psychology and Psychiatry*, *48*(3-4), 329-354.

Feldman, R., & Greenbaum, C. W. (1997). Affect regulation and synchrony in mother—infant play as precursors to the development of symbolic competence. *Infant Mental Health Journal: Official Publication of The World Association for Infant Mental Health*, *18*(1), 4-23.

Feldman, R., Magori-Cohen, R., Galili, G., Singer, M., & Louzoun, Y. (2011). Mother and infant coordinate heart rhythms through episodes of interaction synchrony. *Infant Behavior and Development*, *34*(4), 569-577.

Feng, C., Tian, T., Feng, X., & Luo, Y. J. (2015). Attention modulations on the perception of social hierarchy at distinct temporal stages: An electrophysiological investigation. *Neuropsychologia*, *70*, 21-29.

Ferguson, B., & Lew-Williams, C. (2016). Communicative signals support abstract rule learning by 7-month-old infants. *Scientific reports*, *6*(1), 1-7.

Franchak, J. M., Kretch, K. S., Soska, K. C., & Adolph, K. E. (2011). Head-mounted eye tracking: A new method to describe infant looking. *Child development*, *82*(6), 1738-1750.

Gamliel, H. N., Nevat, M., Probolovski, H. G., Karklinsky, M., Han, S., & Shamay-Tsoory, S. G. (2021). Inter-group conflict affects inter-brain synchrony during synchronized movements. *Neuroimage*, *245*, 118661.

Gaarder, K., Krauskopf, J., Graf, V., Kropfl, W., & Armington, J. C. (1964). Averaged brain activity following saccadic eye movement. *Science*, *146*(3650), 1481-1483.

Gabard-Durnam, L. J., Mendez Leal, A. S., Wilkinson, C. L., & Levin, A. R. (2018). The Harvard Automated Processing Pipeline for Electroencephalography (HAPPE): standardized processing software for developmental and high-artifact data. *Frontiers in neuroscience*, *12*, 97.

Gao, C., Conte, S., Richards, J. E., Xie, W., & Hanayik, T. (2019). The neural sources of N170: Understanding timing of activation in face-selective areas. *Psychophysiology*, *56*(6), e13336.

Georgieva, S., Lester, S., Noreika, V., Yilmaz, M. N., Wass, S., & Leong, V. (2020). Toward the understanding of topographical and spectral signatures of infant movement artifacts in naturalistic EEG. *Frontiers in neuroscience*, *14*, 352.

Giraud, A. L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature neuroscience*, *15*(4), 511-517.

Golumbic, E. M. Z., Ding, N., Bickel, S., Lakatos, P., Schevon, C. A., McKhann, G. M., ... & Schroeder, C. E. (2013). Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party". *Neuron*, *77*(5), 980-991.

Gottlieb, J., Oudeyer, P. Y., Lopes, M., & Baranes, A. (2013). Information-seeking, curiosity, and attention: computational and neural mechanisms. *Trends in cognitive sciences*, *17*(11), 585-593.

Goupil, L., & Proust, J. (2023). Curiosity as a metacognitive feeling. *Cognition*, *231*, 105325.

Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., & Garrod, S. (2013). Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS biology*, *11*(12), e1001752.

Grossmann, T., & Johnson, M. H. (2007). The development of the social brain in human infancy. *European Journal of Neuroscience*, *25*(4), 909-919.

Gugnowska, K., Novembre, G., Kohler, N., Villringer, A., Keller, P. E., & Sammler, D. (2022). Endogenous sources of interbrain synchrony in duetting pianists. *Cerebral Cortex*.

Guy, M. W., Zieber, N., & Richards, J. E. (2016). The cortical development of specialized face processing in infancy. *Child Development*, *87*(5), 1581-1600.

Guy, M. W., Richards, J. E., Tonnsen, B. L., & Roberts, J. E. (2018). Neural correlates of face processing in etiologically-distinct 12-month-old infants at high-risk of autism spectrum disorder. *Developmental Cognitive Neuroscience*, *29*, 61-71.

Haegens, S., & Golumbic, E. Z. (2018). Rhythmic facilitation of sensory processing: A critical review. *Neuroscience & Biobehavioral Reviews*, *86*, 150-165.

Hains, S. M., & Muir, D. W. (1996). Infant sensitivity to adult eye direction. *Child development*, *67*(5), 1940-1951.

Halder, S., Bensch, M., Mellinger, J., Bogdan, M., Kübler, A., Birbaumer, N., & Rosenstiel, W. (2007). Online artifact removal for brain-computer interfaces using support vector machines and blind source separation. *Computational intelligence and neuroscience*, *2007*.

Hale, J., Ward, J. A., Buccheri, F., Oliver, D., & Hamilton, A. F. D. C. (2020). Are you on my wavelength? Interpersonal coordination in dyadic conversations. *Journal of nonverbal behavior*, *44*(1), 63-83.

Hamilton, A. F. D. C. (2016). Gazing at me: the importance of social meaning in understanding direct-gaze cues. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 20150080.

Hamilton, A. F. D. C. (2021). Hyperscanning: beyond the hype. *Neuron*, *109*(3), 404-407.

Hamilton, A. F. D. C., & Lind, F. (2016). Audience effects: what can they tell us about social neuroscience, theory of mind and autism?. *Culture and Brain*, *4*(2), 159-177.

Haresign, I. M., Phillips, E., Whitehorn, M., Noreika, V., Jones, E. J. H., Leong, V., & Wass, S. V. (2021). Automatic classification of ICA components from infant EEG using MARA. *Developmental cognitive neuroscience*, *52*, 101024.

Haresign, I. M., Phillips, E. A. M., Whitehorn, M., Goupil, L., Noreika, V., Leong, V., & Wass, S. V. (2022). Measuring the temporal dynamics of inter-personal neural entrainment in continuous child-adult EEG hyperscanning data. *Developmental cognitive neuroscience*, *54*, 101093.

Hasson, U., & Frith, C. D. (2016). Mirroring and beyond: coupled dynamics as a generalized framework for modelling social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1693), 20150366.

Henry, M. J., & Obleser, J. (2012). Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. *Proceedings of the National Academy of Sciences*, *109*(49), 20095-20100.

Hirsch, J., Zhang, X., Noah, J. A., & Ono, Y. (2017). Frontal temporal and parietal systems synchronize within and across brains during live eye-to-eye contact. *Neuroimage*, *157*, 314-330.

Ho, H. T., Burr, D. C., Alais, D., & Morrone, M. C. (2022). Propagation and update of auditory perceptual priors through alpha and theta rhythms. *European Journal of Neuroscience*, *55*(11-12), 3083-3099.

Hoehl, S., Fairhurst, M., & Schirmer, A. (2021). Interactional synchrony: signals, mechanisms and benefits. *Social Cognitive and Affective Neuroscience*, *16*(1-2), 5-18.

Hoehl, S., & Striano, T. (2008). Neural processing of eye gaze and threat-related emotional facial expressions in infancy. *Child development*, *79*(6), 1752-1760.

Hoehl, S., & Striano, T. (2010). Infants' neural processing of positive emotion and eye gaze. *Social Neuroscience*, *5*(1), 30-39.

Hoehl, S., Wahl, S., Michel, C., & Striano, T. (2012). Effects of eye gaze cues provided by the caregiver compared to a stranger on infants' object processing. *Developmental Cognitive Neuroscience*, *2*(1), 81-89.

Hoehl, S., Wiese, L., & Striano, T. (2008). Young infants' neural processing of objects is affected by eye gaze direction and emotional expression. *PLoS One*, *3*(6), e2389.

Hoffman, K. L., Dragan, M. C., Leonard, T. K., Micheli, C., Montefusco-Siegmund, R., & Valiante, T. A. (2013). Saccades during visual exploration align hippocampal 3–8 Hz rhythms in human and non-human primates. *Frontiers in systems neuroscience*, *7*, 43.

Holleman, G. A., Hooge, I. T., Kemner, C., & Hessels, R. S. (2020). The 'Real-World Approach'and Its Problems: A Critique of the Term Ecological Validity. *Frontiers in Psychology*, *11*, 721.

Holroyd, C. B. (2022). Interbrain synchrony: on wavy ground. *Trends in Neurosciences*.

Hoormann, J., Falkenstein, M., Schwarzenau, P., & Hohnsbein, J. (1998). Methods for the quantification and statistical testing of ERP differences across conditions. *Behavior Research Methods, Instruments, & Computers*, *30*(1), 103-109.

Itier, R. J., Alain, C., Sedore, K., & McIntosh, A. R. (2007). Early face processing specificity: it's in the eyes!. *Journal of cognitive neuroscience*, *19*(11), 1815-1826.

Izuma, K., Saito, D. N., & Sadato, N. (2010). Processing of the incentive for social approval in the ventral striatum during charitable donation. *Journal of cognitive neuroscience*, *22*(4), 621-631.

Izuma, K., Saito, D. N., & Sadato, N. (2010). The roles of the medial prefrontal cortex and striatum in reputation processing. *Social Neuroscience*, *5*(2), 133-147.

Jensen, O., Gips, B., Bergmann, T. O., & Bonnefond, M. (2014). Temporal coding organized by coupled Alpha and gamma oscillations prioritize visual processing. *Trends in neurosciences*, *37*(7), 357-369.

Jones, E. J., & Johnson, M. H. (2017). Early neurocognitive markers of developmental psychopathology. *The Wiley Handbook of Developmental Psychopathology*, 197-214.

Jones, E. J., Goodwin, A., Orekhova, E., Charman, T., Dawson, G., Webb, S. J., & Johnson, M. H. (2020). Infant EEG theta modulation predicts childhood intelligence. *Scientific reports*, *10*(1), 1-10.

Kayhan, E., Matthes, D., Haresign, I. M., Bánki, A., Michel, C., Langeloh, M., ... & Hoehl, S. (2022). DEEP: A dual EEG pipeline for developmental hyperscanning studies. *Developmental cognitive neuroscience*, *54*, 101104.

Kazai, K., & Yagi, A. (2003). Comparison between the lambda response of eye-fixation-related potentials and the P100 component of pattern-reversal visual evoked potentials. *Cognitive, Affective, & Behavioral Neuroscience*, *3*(1), 46-56.

Keitel, C., Obleser, J., Jessen, S., & Henry, M. J. (2021). Frequency-Specific Effects in Infant Electroencephalograms Do Not Require Entrained Neural Oscillations: A Commentary on Köster et al.(2019). *Psychological Science*, 09567976211001317.

Keitel, C., Ruzzoli, M., Dugué, L., Busch, N. A., & Benwell, C. S. (2022). Rhythms in cognition: The evidence revisited. *European Journal of Neuroscience*, *55*(11-12), 2991-3009.

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PloS one*, *7*(5), e36399.

Kingsbury, L., Huang, S., Wang, J., Gu, K., Golshani, P., Wu, Y. E., & Hong, W. (2019). Correlated Neural Activity and Encoding of Behavior across Brains of Socially Interacting Animals. *Cell*.

Kinreich, S., Djalovski, A., Kraus, L., Louzoun, Y., & Feldman, R. (2017). Brain-to-brain synchrony during naturalistic social interactions. *Scientific reports*, *7*(1), 1-12.

Kirkland, A. (2020). An exploration of the neural correlates of turn-taking in spontaneous conversation.

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG Alpha oscillations: the inhibition–timing hypothesis. *Brain research reviews*, *53*(1), 63-88.

Kuhl, P. K., Tsao, F. M., & Liu, H. M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences*, *100*(15), 9096-9101.

Lachaux, J. P., Rodriguez, E., Martinerie, J., & Varela, F. J. (1999). Measuring phase synchrony in brain signals. *Human brain mapping*, *8*(4), 194-208.

Lakatos, P., Gross, J., & Thut, G. (2019). A new unifying account of the roles of neuronal entrainment. *Current Biology*, *29*(18), R890-R905.

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *science*, *320*(5872), 110-113.

Lakatos, P., Musacchia, G., O'Connel, M. N., Falchier, A. Y., Javitt, D. C., & Schroeder, C. E. (2013). The spectrotemporal filter mechanism of auditory selective attention. *Neuron*, *77*(4), 750-761.

Latinus, M., Love, S. A., Rossi, A., Parada, F. J., Huang, L., Conty, L., ... & Puce, A. (2015). Social decisions affect neural activity to perceived dynamic gaze. *Social cognitive and affective neuroscience*, *10*(11), 1557-1567.

Leach, S. C., Morales, S., Bowers, M. E., Buzzell, G. A., Debnath, R., Beall, D., & Fox, N. A. (2020). Adjusting ADJUST: Optimizing the ADJUST algorithm for paediatric data using geodesic nets. *Psychophysiology*, e13566.

Lee, M. D., & Wagenmakers, E. J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge university press.

Lepage, Jean-François, and Hugo Théoret. "EEG evidence for the presence of an action observation–execution matching system in children." *European Journal of Neuroscience* 23.9 (2006): 2505-2510.

Leong, V., Byrne, E., Clackson, K., Georgieva, S., Lam, S., & Wass, S. (2017). Speaker gaze increases information coupling between infant and adult brains. *Proceedings of the National Academy of Sciences*, *114*(50), 13290-13295.

Leong, V., Kalashnikova, M., Burnham, D., & Goswami, U. (2017). The temporal modulation structure of infant-directed speech. *Open Mind*, *1*(2), 78-90.

Leong, V., Noreika, V., Clackson, K., Georgieva, S., Brightman, L., Nutbrown, R., . . . Wass, S. (2019). Mother-infant interpersonal neural entrainment predicts infants' social learning.

Lester, B. M., Hoffman, J., & Brazelton, T. B. (1985). The rhythmic structure of mother-infant interaction in term and preterm infants. *Child development*, 15-27.

Lin, W. M., Oetringer, D. A., Bakker-Marshall, I., Emmerzaal, J., Wilsch, A., ElShafei, H. A., ... & Haegens, S. (2022). No behavioural evidence for rhythmic facilitation of perceptual discrimination. *European Journal of Neuroscience*, *55*(11-12), 3352-3364.

Lindenberger, U., Li, S. C., Gruber, W., & Müller, V. (2009). Brains swinging in concert: cortical phase synchronization while playing guitar. *BMC neuroscience*, *10*(1), 1-12.

Liu, D., Liu, S., Liu, X., Zhang, C., Li, A., Jin, C., . . . Zhang, X. (2018). Interactive Brain Activity: Review and Progress on EEG-Based Hyperscanning in Social Interactions. *Frontiers in Psychology, 9*, 1862.

Lobier, M., Siebenhühner, F., Palva, S., & Palva, J. M. (2014). Phase transfer entropy: a novel phase-based measure for directed entrainment in networks coupled by oscillatory interactions. *Neuroimage*, *85*, 853-872.

Loomis, J. M., Kelly, J. W., Pusch, M., Bailenson, J. N., & Beall, A. C. (2008). Psychophysics of perceiving eye-gaze and head direction with peripheral vision: Implications for the dynamics of eye-gaze behavior. *Perception*, *37*(9), 1443-1457.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *nature*, *412*(6843), 150-157.

López Pérez, D., Leonardi, G., Niedźwiecka, A., Radkowska, A., Rączaszek-Leonardi, J., & Tomalski, P. (2017). Combining recurrence analysis and automatic movement extraction from video recordings to study behavioral coupling in face-to-face parent-child interactions. *Frontiers in psychology*, *8*, 2228.

Luft, C., Zioga, I., Giannopoulos, A., Di Bona, G., Civilini, A., Latora, V., & Mareschal, I. (2021). Social synchronisation of brain activity by eye-contact.

Luck, S. J. (2014). An introduction to the event-related potential technique. MIT press.

Luck, S. J., Stewart, A. X., Simmons, A. M., & Rhemtulla, M. (2021). Standardized measurement error: A universal metric of data quality for averaged event-related potentials. *Psychophysiology*, *58*(6), e13793.

Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., ... & Grundmann, M. (2019). Mediapipe: A framework for building perception pipelines. *arXiv preprint arXiv:1906.08172*.

Makeig, S., Bell, A. J., Jung, T. P., & Sejnowski, T. J. (1996). Independent component analysis of electroencephalographic data. In *Advances in neural information processing systems*(pp. 145-151).

Makeig, S., Debener, S., Onton, J., & Delorme, A. (2004). Mining event-related brain dynamics. *Trends in cognitive sciences*, *8*(5), 204-210.

Mandel, A., Bourguignon, M., Parkkonen, L., & Hari, R. (2016). Sensorimotor activation related to speaker vs. listener role during natural conversation. *Neuroscience letters*, *614*, 99-104.

Mansour, H., & Kuhn, G. (2019). Studying “natural” eye movements in an “unnatural” social environment: The influence of social activity, framing, and sub-clinical traits on gaze aversion. *Quarterly Journal of Experimental Psychology*, *72*(8), 1913-1925.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of neuroscience methods*, *164*(1), 177-190.

Marshall, P. J., Bar-Haim, Y., & Fox, N. A. (2002). Development of the EEG from 5 months to 4 years of age. *Clinical Neurophysiology*, *113*(8), 1199-1208.

Marshall, P. J., Young, T., & Meltzoff, A. N. (2011). Neural correlates of action observation and execution in 14-month-old infants: An event-related EEG desynchronization study. *Developmental science*, *14*(3), 474-480.

Mathewson, K. E., Fabiani, M., Gratton, G., Beck, D. M., & Lleras, A. (2010). Rescuing stimuli from invisibility: Inducing a momentary release from visual masking with pre-target entrainment. *Cognition*, *115*(1), 186-191.

Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M., & Ro, T. (2009). To see or not to see: prestimulus α phase predicts visual awareness. *Journal of Neuroscience*, *29*(9), 2725-2732.

Mathewson, K. E., Lleras, A., Beck, D. M., Fabiani, M., Ro, T., & Gratton, G. (2011). Pulsed out of awareness: EEG Alpha oscillations represent a pulsed-inhibition of ongoing cortical processing. *Frontiers in psychology*, *2*, 99.

Mathewson, K. E., Prudhomme, C., Fabiani, M., Beck, D. M., Lleras, A., & Gratton, G. (2012). Making waves in the stream of consciousness: entraining oscillations in EEG Alpha and fluctuations in visual awareness with rhythmic visual stimulation. *Journal of cognitive neuroscience*, *24*(12), 2321-2333.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica*, *22*(3), 276-282.

Michel, R., Dugué, L., & Busch, N. A. (2022). Distinct contributions of alpha and theta rhythms to perceptual and attentional sampling. *European Journal of Neuroscience*, *55*(11-12), 3025-3039.

Moreau, Q., & Dumas, G. (2021). Beyond "correlation vs. causation": multi-brain neuroscience needs explanation. *Trends Cogn. Sci. Published online March*, *19*, 2021.

Morrow, A., & Samaha, J. (2022). No evidence for a single oscillator underlying discrete visual percepts. *European Journal of Neuroscience*, *55*(11-12), 3054-3066.

Mullen, T. (2012). CleanLine EEGLAB plugin. *San Diego, CA: Neuroimaging Informatics Tools and Resources Clearinghouse (NITRC)*.

Murray, L., De Pascalis, L., Bozicevic, L., Hawkins, L., Sclafani, V., & Ferrari, P. F. (2016). The functional architecture of mother-infant communication, and the development of infant social expressiveness in the first two months. *Scientific Reports*, *6*(1), 1-9

Muthukumaraswamy, S. D., & Singh, K. D. (2011). A cautionary note on the interpretation of phase-locking estimates with concurrent changes in power. *Clinical neurophysiology: official journal of the International Federation of Clinical Neurophysiology*, *122*(11), 2324-2325.

Nguyen, T., Schleihauf, H., Kayhan, E., Matthes, D., Vrtička, P., & Hoehl, S. (2020). The effects of interaction quality on neural synchrony during mother-child problem solving. *cortex*, *124*, 235-249.

Nguyen, T., Schleihauf, H., Kayhan, E., Matthes, D., Vrtička, P., & Hoehl, S. (2021). Neural synchrony in mother–child conversation: Exploring the role of conversation patterns. *Social Cognitive and Affective Neuroscience*, *16*(1-2), 93-102.

Noah, J. A., Zhang, X., Dravida, S., Ono, Y., Naples, A., McPartland, J. C., & Hirsch, J. (2020). Real-time eye-to-eye contact is associated with cross-brain neural coupling in angular gyrus. *Frontiers in human neuroscience*, *14*, 19.

Noreika, V., Georgieva, S., Wass, S., & Leong, V. (2020). 14 challenges and their solutions for conducting social neuroscience and longitudinal EEG research with infants. *Infant Behavior and Development*, *58*, 101393.

Novembre, G., & Iannetti, G. D. (2021). Hyperscanning alone cannot prove causality. Multibrain stimulation can. *Trends in Cognitive Sciences*, *25*(2), 96-99.

Obleser, J., & Kayser, C. (2019). Neural entrainment and attentional selection in the listening brain. *Trends in cognitive sciences*, *23*(11), 913-926.

Ono, K., Hashimoto, J., & Sasaoka, T. (2022). Intertap interval dependence of the subdivision effect in auditory-synchronised tapping. *European Journal of Neuroscience*, *55*(11-12), 3391-3401.

Orekhova, E. V., Elsabbagh, M., Jones, E. J., Dawson, G., Charman, T., & Johnson, M. H. (2014). EEG hyper-connectivity in high-risk infants is associated with later autism. *Journal of neurodevelopmental disorders*, *6*(1), 1-11.

Orekhova, E. V., Stroganova, T. A., Posikera, I. N., & Elam, M. (2006). EEG theta rhythm in infants and preschool children. *Clinical Neurophysiology*, *117*(5), 1047-1062.

Oudeyer, P. Y., & Smith, L. B. (2016). How evolution may work through curiosity-driven developmental process. *Topics in Cognitive Science*, *8*(2), 492-502.

Parise, E., Reid, V. M., Stets, M., & Striano, T. (2008). Direct eye contact influences the neural processing of objects in 5-month-old infants. *Social Neuroscience*, *3*(2), 141-150.

Pan, Y., Novembre, G., & Olsson, A. (2021). The interpersonal neuroscience of social learning. *Perspectives on Psychological Science*, 17456916211008429.

Peykarjou, S., & Hoehl, S. (2013). Three-month-olds' brain responses to upright and inverted faces and cars. *Developmental Neuropsychology*, *38*(4), 272-280.

Pérez, A., Carreiras, M., & Duñabeitia, J. A. (2017). Brain-to-brain entrainment: EEG interbrain synchronization while speaking and listening. *Scientific reports*, *7*(1), 1-12.

Piazza, C., Cantiani, C., Miyakoshi, M., Riva, V., Molteni, M., Reni, G., & Makeig, S. (2020). EEG effective source projections are more bilaterally symmetric in infants than in adults. *Frontiers in Human Neuroscience*, *14*.

Piazza, E. A., Cohen, A., Trach, J., & Lew-Williams, C. (2021). Neural synchrony predicts children's learning of novel words. *Cognition*, *214*, 104752.

Piazza, E. A., Hasenfratz, L., Hasson, U., & Lew-Williams, C. (2020). Infant and adult brains are coupled to the dynamics of natural communication. *Psychological Science*, *31*(1), 6-17.

Pinti, P., Siddiqui, M. F., Levy, A. D., Jones, E. J. H., & Tachtsidis, I. (2021). An analysis framework for the integration of broadband NIRS and EEG to assess neurovascular and neurometabolic coupling. *Scientific reports*, *11*(1), 1-20.

Pérez, A., Carreiras, M., & Duñabeitia, J. A. (2017). Brain-to-brain entrainment: EEG interbrain synchronization while speaking and listening. *Scientific reports*, *7*(1), 1-12.

Pérez, A., Dumas, G., Karadag, M., & Duñabeitia, J. A. (2019). Differential brain-to-brain entrainment while speaking and listening in native and foreign languages. *Cortex*, *111*, 303-315.

Phillips, E., Goupil, L., Haresign, I. M., Bruce-Gardyne, E., Csolsim, F. A., Whitehorn, M., ... & Wass, S. (2021). Proactive or reactive? Neural oscillatory insight into the leader-follower dynamics of early infant-caregiver interaction.

Plöchl, M., Fiebelkorn, I., Kastner, S., & Obleser, J. (2022). Attentional sampling of visual and auditory objects is captured by theta-modulated neural activity. *European Journal of Neuroscience*, *55*(11-12), 3067-3082.

Plöchl, M., Ossandón, J. P., & König, P. (2012). Combining EEG and eye tracking: identification, characterization, and correction of eye movement artifacts in electroencephalographic data. *Frontiers in human neuroscience*, *6*, 278.

Poeppel, D., & Assaneo, M. F. (2020). Speech rhythms and their neural foundations. *Nature reviews neuroscience*, *21*(6), 322-334.

Poli, F., Serino, G., Mars, R. B., & Hunnius, S. (2020). Infants tailor their attention to maximize learning. *Science advances*, *6*(39), eabb5053.

Pönkänen, L. M., Alhoniemi, A., Leppänen, J. M., & Hietanen, J. K. (2011). Does it make a difference if I have an eye contact with you or with your picture? An ERP study. *Social cognitive and affective neuroscience*, *6*(4), 486-494.

Pönkänen, L. M., Peltola, M. J., & Hietanen, J. K. (2011). The observer observed: Frontal EEG asymmetry and autonomic responses differentiate between another person's direct and averted gaze when the face is seen live. *International Journal of Psychophysiology*, *82*(2), 180-187.

Quadrelli, E., Conte, S., Macchi Cassia, V., & Turati, C. (2019). Emotion in motion: Facial dynamics affect infants' neural processing of emotions. *Developmental psychobiology*, *61*(6), 843-858

Rassili, O., & Ordin, M. (2022). The effect of regular rhythm on the perception of linguistic and non-linguistic auditory input. *European Journal of Neuroscience*, *55*(11-12), 3365-3372.

Redcay, E., & Schilbach, L. (2019). Using second-person neuroscience to elucidate the mechanisms of social interaction. *Nature Reviews Neuroscience*, *20*(8), 495-505.

Redcay, E., & Warnell, K. R. (2018). A social-interactive neuroscience approach to understanding the developing brain. *Advances in child development and behavior*, *54*, 1-44.

Reindl, V., Gerloff, C., Scharke, W., & Konrad, K. (2018). Brain-to-brain synchrony in parent-child dyads and the relationship with emotion regulation revealed by fNIRS-based hyperscanning. *NeuroImage, 178*, 493-502.

Reindl, V., Wass, S., Leong, V., Scharke, W., Wistuba, S., Wirth, C. L., ... & Gerloff, C. (2021). Synchrony of mind and body are distinct in mother-child dyads. *bioRxiv*.

Reinero, D. A., Dikker, S., & Van Bavel, J. J. (2021). Inter-brain synchrony in teams predicts collective performance. *Social cognitive and affective neuroscience*, *16*(1-2), 43-57.

Richards, J. E. (2001). Cortical indexes of saccade planning following covert orienting in 20-week-old infants. *Infancy*, *2*(2), 135-157.

Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). The art of conversation is coordination. *Psychological science*, *18*(5), 407-413.

Risko, E.F., Richardson, D.C., Kingstone, A., 2016. Breaking the Fourth Wall of Cognitive Science: Real-World Social Attention and the Dual Function of Gaze. *Curr. Dir. Psychol. Sci*. 25, 70–74. https://doi.org/10.1177/0963721415617806

Robertson, S. S., Guckenheimer, J., Masnick, A. M., & Bacher, L. F. (2004). The dynamics of infant visual foraging.

Rogoff, B. (1990). Apprenticeship in thinking: Cognitive development in social context. Oxford university press.

Ronconi, L., Oosterhof, N. N., Bonmassar, C., & Melcher, D. (2017). Multiple oscillatory rhythms determine the temporal organization of perception. *Proceedings of the National Academy of Sciences*, *114*(51), 13435-13440.

Rossi, A., Parada, F. J., Kolchinsky, A., & Puce, A. (2014). Neural correlates of apparent motion perception of impoverished facial stimuli: a comparison of ERP and ERSP activity. *NeuroImage*, *98*, 442-459.

Rousselet, G. A., Husk, J. S., Bennett, P. J., & Sekuler, A. B. (2007). Single-trial EEG dynamics of object and face visual processing. *Neuroimage*, *36*(3), 843-862.

Rutledge, D. N., & Bouveresse, D. J. R. (2013). Independent components analysis with the JADE algorithm. *TrAC Trends in Analytical Chemistry*, *50*, 22-32.

Ruzzoli, M., Torralba, M., Fernández, L. M., & Soto-Faraco, S. (2019). The relevance of Alpha phase in human perception. *Cortex*, *120*, 249-268.

Samaha, J., & Postle, B. R. (2015). The speed of alpha-band oscillations predicts the temporal resolution of visual perception. *Current Biology*, *25*(22), 2985-2990.

Sameshima, K., & Baccala, L. A. (Eds.). (2014). *Methods in brain entrainment inference through multivariate time series analysis*. CRC press.

Sänger, J., Müller, V., & Lindenberger, U. (2013). Directionality in hyperbrain networks discriminates between leaders and followers in guitar duets. *Frontiers in human neuroscience*, *7*, 234.

Santamaria, L., Noreika, V., Georgieva, S., Clackson, K., Wass, S., & Leong, V. (2020). Emotional valence modulates the topology of the parent-infant inter-brain network. *NeuroImage*, *207*, 116341.

Sauseng, P., Klimesch, W., Gruber, W. R., Hanslmayr, S., Freunberger, R., & Doppelmayr, M. (2007). Are event-related potential components generated by phase resetting of brain oscillations? A critical discussion. *Neuroscience*, *146*(4), 1435-1444.

Schapker, H., Breithaupt, T., Shuranova, Z., Burmistrov, Y., & Cooper, R. L. (2002). Heart and ventilatory measures in crayfish during environmental disturbances and social interactions. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, *131*(2), 397-407.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience 1. *Behavioral and brain sciences*, *36*(4), 393-414.

Schiller, P. H., Slocum, W. M., Carvey, C., & Tolias, A. S. (2004). Are express saccades generated under natural viewing conditions?. *European Journal of Neuroscience*, *20*(9), 2467-2473.

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in neurosciences*, *32*(1), 9-18.

Senju, A., & Johnson, M. H. (2009). The eye contact effect: mechanisms and development. *Trends in cognitive sciences*, *13*(3), 127-134.

Seth, A. K., Barrett, A. B., & Barnett, L. (2015). Granger causality analysis in neuroscience and neuroimaging. *Journal of Neuroscience*, *35*(8), 3293-3297.

Shamay-Tsoory, S. G., Saporta, N., Marton-Alper, I. Z., & Gvirts, H. Z. (2019). Herding brains: a core neural mechanism for social alignment. *Trends in cognitive sciences*, *23*(3), 174-186.

Simony, E., Honey, C. J., Chen, J., Lositsky, O., Yeshurun, Y., Wiesel, A., & Hasson, U. (2016). Dynamic reconfiguration of the default mode network during narrative comprehension. *Nature communications, 7*, 12141.

Southgate, V., Chevallier, C., & Csibra, G. (2010). Seventeen-month-olds appeal to false beliefs to interpret others' referential communication. *Developmental science*, *13*(6), 907-912.

Stefan, A. M., Gronau, Q. F., Schönbrodt, F. D., & Wagenmakers, E. J. (2019). A tutorial on Bayes Factor Design Analysis using an informed prior. *Behavior research methods*, *51*(3), 1042-1058.

Stephani, T., Driller, K. K., Dimigen, O., & Sommer, W. (2020). Eye contact in active and passive viewing: Event-related brain potential evidence from a combined eye tracking and EEG study. *Neuropsychologia*, *143*, 107478.

Stern, D. N. (1971). A Micro-Analysis of Mother-Infant Interaction. Behavior Regulating Social Contact Between a Mother and her 3 1/2 Month-Old Twins. *Journal of the american academy of child psychiatry*, *10*(3), 501-517.

Sun, Y., Michalareas, G., & Poeppel, D. (2022). The impact of phase entrainment on auditory detection is highly variable: Revisiting a key finding. *European Journal of Neuroscience*, *55*(11-12), 3373-3390.

Suppanen, E., Huotilainen, M., & Ylinen, S. (2019). Rhythmic structure facilitates learning from auditory input in newborn infants. *Infant Behavior and Development*, *57*, 101346.

Symons, L. A., Hains, S. M., & Muir, D. W. (1998). Look at me: Five-month-old infants' sensitivity to very small deviations in eye-gaze during social interactions. *Infant Behavior and Development*, *21*(3), 531-536.

Tass P, Rosenblum MG, Weule J, Kurths J, Pikovsky A, Volkmann J, Schnitzler A, and Freund HJ. Detection of n:m phase locking from noisy data: application to magnetoencephalography. Phys Rev Lett 81: 3291–3294, 1998.

Taylor, M. J., Batty, M., & Itier, R. J. (2004). The faces of development: a review of early face processing over childhood. *Journal of cognitive neuroscience*, *16*(8), 1426-1442.

Taylor, M. J., Edmonds, G. E., McCarthy, G., & Allison, T. (2001). Eyes first! Eye processing develops before face processing in children. *Neuroreport*, *12*(8), 1671-1676.

Thickbroom, G. W., Knezevic, W., Carroll, W. M., & Mastaglia, F. L. (1991). Saccade onset and offset lambda waves: relation to pattern movement visually evoked potentials. *Brain research*, *551*(1-2), 150-156.

Thaut, M. H., McIntosh, G. C., & Hoemberg, V. (2015). Neurobiological foundations of neurologic music therapy: rhythmic entrainment and the motor system. *Frontiers in psychology*, *5*, 1185.

Thut, G., Miniussi, C., & Gross, J. (2012). The functional importance of rhythmic activity in the brain. *Current Biology*, *22*(16), R658-R663.

Tort, A. B., Komorowski, R., Eichenbaum, H., & Kopell, N. (2010). Measuring phase-amplitude coupling between neuronal oscillations of different frequencies. *Journal of neurophysiology*, *104*(2), 1195-1210.

VanRullen, R. (2016). Perceptual cycles. *Trends in cognitive sciences*, *20*(10), 723-735.

Van der Velde, B., White, T., & Kemner, C. (2021). The emergence of a theta social brain network during infancy. *NeuroImage*, *240*, 118298.

Wang, X., Chen, Y., & Ding, M. (2008). Estimating Granger causality after stimulus onset: a cautionary note. *Neuroimage*, *41*(3), 767-776.

Wass, S., & Goupil, L. (2022). Studying the developing brain in real-world contexts: moving from castles in the air to castles on the ground.

Wass, S. V., Noreika, V., Georgieva, S., Clackson, K., Brightman, L., Nutbrown, R., ... & Leong, V. (2018). Parental neural responsivity to infants' visual attention: How mature brains influence immature brains during social interaction. *PLoS biology*, *16*(12), e2006328.

Wass, S. V., Whitehorn, M., Haresign, I. M., Phillips, E., & Leong, V. (2020). Interpersonal neural entrainment during early social interaction. *Trends in cognitive sciences*, *24*(4), 329-342.

Watanabe, S., Kakigi, R., & Puce, A. (2001). Occipitotemporal activity elicited by viewing eye movements: a magnetoencephalographic study. *Neuroimage*, *13*(2), 351-363.

Watanabe, S., Miki, K., & Kakigi, R. (2002). Gaze direction affects face perception in humans. *Neuroscience letters*, *325*(3), 163-166.

Webb, S. J., Bernier, R., Henderson, H. A., Johnson, M. H., Jones, E. J., Lerner, M. D., ... & Westerfield, M. (2015). Guidelines and best practices for electrophysiological data collection, analysis and reporting in autism. *Journal of autism and developmental disorders*, *45*(2), 425-443.

Werker, J. F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., & Amano, S. (2007). Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition*, *103*(1), 147-162.

Werker, J. F., & Yeung, H. H. (2005). Infant speech perception bootstraps word learning. *Trends in cognitive sciences*, *9*(11), 519-527.

Widmann, A., & Schröger, E. (2012). Filter effects and filter artifacts in the analysis of electrophysiological data. *Frontiers in psychology*, *3*, 233.

Winkler, I., Brandl, S., Horn, F., Waldburger, E., Allefeld, C., & Tangermann, M. (2014). Robust artifactual independent component classification for BCI practitioners. *Journal of neural engineering*, *11*(3), 035013.

Winkler, I., Debener, S., Müller, K. R., & Tangermann, M. (2015, August). On the influence of high-pass filtering on ICA-based artifact reduction in EEG-ERP. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 4101-4105). IEEE.

Winkler, I., Haufe, S., & Tangermann, M. (2011). Automatic classification of artifactual ICA-components for artifact removal in EEG signals. *Behavioural and brain functions*, *7*(1), 30.

Wutz, A., Melcher, D., & Samaha, J. (2018). Frequency modulation of neural oscillations according to visual task demands. *Proceedings of the National Academy of Sciences*, *115*(6), 1346-1351.

Van Diepen, R. M., Cohen, M. X., Denys, D., & Mazaheri, A. (2015). Attention and temporal expectations modulate power, not phase, of ongoing Alpha oscillations. *Journal of cognitive neuroscience*, *27*(8), 1573-1586.

VanRullen, R. (2016). Perceptual cycles. *Trends in cognitive sciences*, *20*(10), 723-735.

Vecera, S. P., & Johnson, M. H. (1995). Gaze detection and the cortical processing of faces: Evidence from infants and adults. *Visual cognition*, *2*(1), 59-87.

Yu, C., & Smith, L. B. (2013). Joint attention without gaze following: Human infants and their parents coordinate visual attention to objects through eye-hand coordination. *PloS one*, *8*(11), e79659.

Yu, C. and L.B. Smith, The social origins of sustained attention in one-year-old human infants. Current Biology, 2016. 26(9): p. 1235-1240.

Yuval-Greenberg, S., Tomer, O., Keren, A. S., Nelken, I., & Deouell, L. Y. (2008). Transient induced gamma-band response in EEG as a manifestation of miniature saccades. *Neuron, 58*(3), 429-441.

Xie, W., Mallin, B. M., & Richards, J. E. (2018). Development of infant sustained attention and its relation to EEG oscillations: an EEG and cortical source analysis study. *Developmental Science*, *21*(3), e12562.

Xie, W., & Richards, J. E. (2016). Effects of interstimulus intervals on behavioral, heart rate, and event-related potential indices of infant engagement and sustained attention. *Psychophysiology*, *53*(8), 1128-1142.

Yun, K., Watanabe, K., & Shimojo, S. (2012). Interpersonal body and neural synchronization as a marker of implicit social interaction. *Scientific reports*, *2*(1), 1-8.

Zhang, W., & Yartsev, M. M. (2019). Correlated Neural Activity across the Brains of Socially Interacting Bats. *Cell*.