

14 challenges and their solutions for conducting social neuroscience and longitudinal EEG research with infants

Valdas Noreika^{1,*}, Stanimira Georgieva¹, Sam Wass³, Victoria Leong^{1,2}

¹ Department of Psychology, University of Cambridge, Cambridge, UK

² Division of Psychology, Nanyang Technological University, Singapore

³ School of Psychology, University of East London, London, UK

* Corresponding author. E-mail: vn261@cam.ac.uk Address: Department of Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, United Kingdom

Abstract

The use of electroencephalography (EEG) to study infant brain development is a growing trend. In addition to classical longitudinal designs that study the development of the neural, cognitive and behavioural function, new areas of EEG application are emerging, such as novel social neuroscience paradigms using dual infant-adult EEG recordings. However, most of the experimental designs, analysis methods, as well as EEG hardware were originally developed for single-person adult research. When applied to the study of infant development, adult-based solutions often pose unique problems that may go unrecognised. Here, we identify 14 challenges that infant EEG researchers may encounter when designing new experiments, collecting data, and conducting data analysis. Challenges related to the experimental design are: (1) small sample size and data attrition, and (2) varying arousal in younger infants. Challenges related to data acquisition are: (3) determining the optimal location for reference and ground electrodes, (4) control of impedance when testing with the high-density sponge electrode nets, (5) poor fit of standard EEG caps to the varying infant head shapes, and (6) ensuring a high degree of temporal synchronisation between amplifiers and recording devices during dual-EEG acquisition. Challenges related to the analysis of longitudinal and social neuroscience datasets are: (7) developmental changes in head anatomy, (8) prevalence and diversity of infant myogenic artefacts, (9) a lack of stereotypical topography of eye movements needed for the ICA-based data cleaning, (10) and relatively high inter-individual variability of EEG responses in younger cohorts. Additional challenges for the analysis of dual EEG data are: (11) developmental shifts in canonical EEG rhythms and difficulties in differentiating true inter-personal synchrony from spurious synchrony due to (12) common intrinsic properties of the signal and (13) shared external perturbation. Finally, (14) there is a lack of test-retest reliability studies of infant EEG. We describe each

of these challenges and suggest possible solutions. While we focus specifically on the social neuroscience and longitudinal research, many of the issues we raise are relevant for all fields of infant EEG research.

Keywords: infant; electroencephalography (EEG); longitudinal design; methodology; social neuroscience.

Introduction

The use of electroencephalography (EEG) to study infant brain development has undergone a slow but steady increase in recent decades. A PubMed search using the words “newborn / neonate / neonatal / baby / infant” and “EEG / electroencephalography / ERP / event-related potential” in the Title or Abstract returned 36 studies published in 1998, 90 studies published in 2008, and 197 studies published in 2018. The rich history of infant EEG research dates back to the 1930s, when J. Roy Smith (1938a-c, 1939) carried out seminal studies on the development of awake and sleep EEG rhythms, including alpha waves, during normal infancy and childhood. Research of infant EEG consolidated in the 1960s and 1970s when basic research into normal infant EEG development (e.g. Hagne, 1968, 1972) extended to examine EEG of premature infants (e.g. Parmelee Jr et al., 1968; Weitzman & Graziani, 1968). While in the subsequent decades infant EEG studies were mostly clinical, e.g. aiming to assess sources of epileptic seizures (Clancy & Legido, 1991) or the effects of exposure to addictive substances (Legido et al., 1992), more recent studies have started to address psychological and cognitive neuroscience topics, such as language development (Kühn-Popp et al., 2016) and gaze monitoring (Hoehl et al., 2014).

EEG research in developmental social neuroscience has used a range of paradigms (see Figure 1). Historically the majority have been ERP studies that typically examine infants' neural responses to static pictures of faces. Well-studied contrasts include differential neural responses to face vs non-face stimuli (e.g. Halit et al., 2004); to faces either gazing directly to the camera, or with gaze averted (e.g. Farroni et al., 2002); and to faces with different emotional expressions (e.g. Nelson & De Haan, 1996).

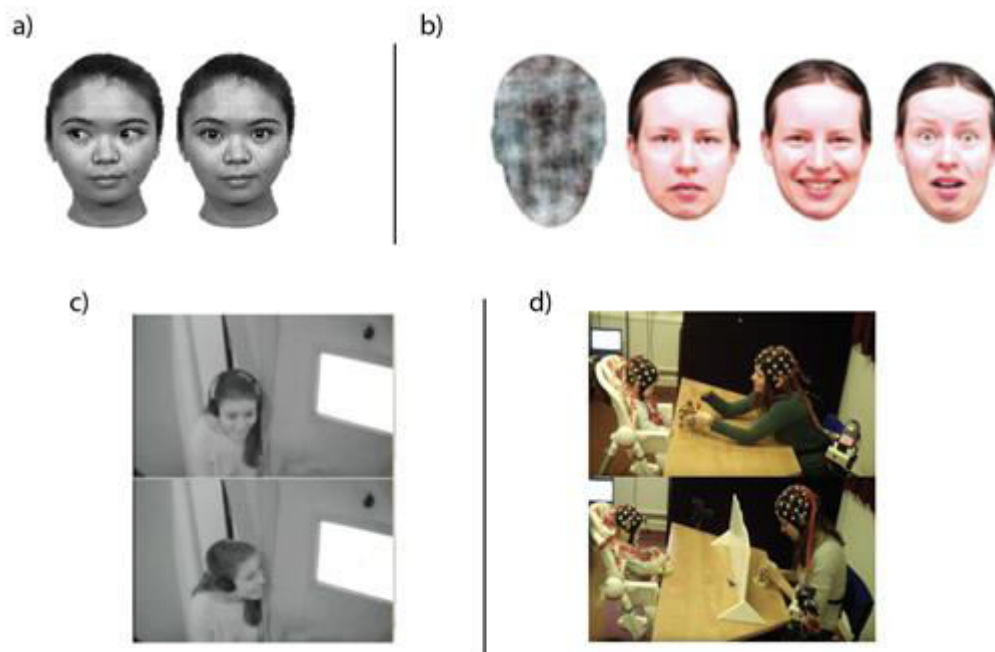


Figure 1 | Illustrating the range of EEG paradigms used in infant social neuroscience research. a) static images showing direct and averted gaze (reproduced from Farroni et al. (2002), Copyright © 2002, The National Academy of Sciences); b) static images showing different emotional expressions (reproduced with permission from Leppänen et al. (2011)); c) interactions featuring a live experimenter and screen-based stimuli (reproduced with permission from Hoehl et al. (2014)); d) live, tabletop interactions with dual EEG recording, reproduced from Wass et al. (2018).

While the strength of these paradigms is their high reproducibility, a limitation is that the paradigms lack ecological validity. Recognising this, other studies have examined differential ERP responses to direct and averted gaze – but including live actors, with onsets and offsets controlled by showing real faces through a computer-controlled electronic shutter (Pönkänen et al., 2011). Other studies have presented dynamic videos of actors talking to the camera (Jones et al., 2015), or real live interactions (Reid et al., 2011), and analysed them using Fourier-transform-based analyses to examine topographic changes in activity across frequencies between social and non-social conditions. Other studies have used conditions in which a live adult experimenter gazes first at the infant’s face and then to a computer-displayed novel object, while the infant’s neural response to the object is measured (Striano et al., 2006; Hoehl et al., 2014).

Although more ecologically valid, one further limitation of these studies is that, by only recording from a single participant at once, they are unable to capture the interpersonal dynamics of how information is shared between two interacting partners during social interaction (Schilbach et al., 2013). To address this, researchers have recently started to build on methodological innovations in adult research (Lachat & George, 2012; Liu et al., 2018) that allows the recording of dual EEG from infants and their adult social partners concurrently *during* an interaction. Types of interaction studied include semi-structured interactions with an unfamiliar adult (Leong et al., 2017), and tabletop play or other cooperative activities with a parent (Wass et al., 2018). Analyses typically concentrate on measuring either Granger-causal relationships between the infant’s and adult’s brain activity and attention (Granger, 1969), or concurrent synchrony (Lachaux et al., 1999).

One problem common for infancy researchers is that most of the hardware and analysis methods, as well as EEG hardware, were originally developed for adult research. When applied to the study of infant development, adult-based solutions often pose unique problems that may go unrecognised. In addition, some of the more newly developed techniques, such as dual EEG recording, pose specific methodological challenges that have been little discussed hitherto. Here, we identify 14 challenges that infant EEG researchers may encounter when designing new experiments, collecting data and conducting data analysis. Several of the methodological problems and possible solutions that we describe in the following are specific to the dual EEG analyses, whereas others are relevant for all areas of infant EEG research, including longitudinal and cross-sectional designs. Given that our aim is to illustrate the identified problems and suggest possible solutions rather than provide an extensive review of infant EEG research, the citation of the infant EEG literature will be selective. The key terms used in the following sections are introduced in Table 1, including suggested references for more information.

Table 1. Key terms related to the acquisition and analysis of EEG data

Concept	Description
Dual EEG	Concurrent acquisition of electroencephalography signals from two different individuals, typically for the purpose of obtaining measures of interpersonal synchronisation. This method is also known as hyperscanning. Babiloni & Astolfi, 2014.
Electrode placement	The positioning of EEG electrodes on the scalp, based on pre-determined anatomical landmarks, e.g. distance from inion. Electrode placement can be carried out manually by calculating the location of each electrode, or – more frequently – by placing an age-appropriate EEG cap/net with fixed electrode

	positions. Fisch & Spehlmann, 1999.
Event-related potentials	Stereotypical patterns of EEG voltage fluctuation that are time-locked to specific sensory stimuli or motor responses. Due to the high amplitude of stimulus-unrelated background EEG activity, the repetition of tens or hundreds of trials per condition is usually required for a reliable estimation of event-related potentials. Luck, 2014.
Granger causality	A directional measure of how one time-series (x) forward-predicts another time series (y), after taking into account the self-predictive past values of y. Granger, 1969.
Ground electrode	An electrode used for common-mode rejection: it takes signals read by all electrodes, such as power line noise (50/60 Hz), and cancels them out of the EEG recordings. In adult studies, this is often placed on the forehead or cheek. Nunez & Srinivasan, 2006.
Impedance	The degree of opposition to alternating current (AC) flow, arising from the combined effects of resistance and reactance. Generally, the higher the impedance of an EEG electrode on the scalp, the smaller the amplitude of the signal recorded by that electrode. The electrode contact impedance is usually kept below 10 k Ω , although much higher impedance values (<100 k Ω) are acceptable when using the high-density Philips-EGI Geodesic Sensor nets. Nunez & Srinivasan, 2006.
Phase coupling	A measure of the association between the on-going oscillatory phase of two EEG time series (e.g. recorded by different electrodes or different individuals). Phase coupling may be assessed by a wide range of metrics, such as the phase-locking value (PLV), phase lag index, mutual information, etc. Cohen, 2014.

Reference electrode	The EEG signal recorded at each electrode is a difference between the signal measured by the electrode itself and another “reference electrode”. In a unipolar montage, the reference electrode is typically placed on the vertex, nose or mastoid. In a bipolar montage, there is a separate reference electrode for each data channel. Nunez & Srinivasan, 2006.
Resting-state	A neuroimaging paradigm in which participants are instructed to relax and stay still without engaging in an active task. In infants, resting-state recordings simulate periods of quiet wakefulness with minimal motor activity. Anderson & Perone, 2018.
Rhythms	The EEG signal is partially composed of neuronal oscillations which can be functionally separated into different frequency ranges, often referred to as EEG rhythms or frequency bands. One classic example is the alpha rhythm (8-12 Hz in adults and 6-9 Hz in infants), which peaks in amplitude over occipital channels during eyes-closed relaxed wakefulness. Buzsaki, 2006.
Source localisation	Identification of the source(s) within the brain from which an EEG signal (or an element of the signal) originates. It involves the prediction of EEG potentials from a model of neural sources (termed as the forward problem) and the estimation of the neural sources from EEG potentials (termed as the inverse problem). Michel & Brunet, 2019.
Spectral power	The squared amplitude of the neural signal typically integrated over a range of signal frequencies. Commonly computed using a Fast Fourier Transform (FFT). Spectral power is usually measured in continuous EEG datasets. Cohen, 2014.

1. Experimental design

1.1 Sample size and data attrition for infant neuroimaging studies

Problem:

Compared to adult studies, infant neuroimaging studies typically suffer from a higher rate of data attrition due to infant non-compliance and signal contamination by motion-related artefacts (Bell & Cuevas, 2012). Whereas adults may be instructed to avoid movement, this strategy is ineffective for paediatric populations. A proportion of infants who attend sessions will either be too fussy or irritable for the EEG cap to be applied in the first place - or will become too upset when the cap is applied, meaning that it immediately has to be removed before data can be collected (see De Haan, 2013, pp 17-18). Furthermore, for some infants, their amount of movement artefact-free data may not meet minimum requirements for data analysis. For example, in resting state paradigms, it is recommended that at least 40-60s of usable data per participant is required for consistent and reliable results (Salinsky et al, 1991; Lund et al, 1995). Therefore, infants who do not contribute to this minimum amount of movement-free data will have to be excluded entirely from analyses. This results in a lower fraction of usable data per participant, and potentially fewer usable datasets (e.g. if minimum requirements for data analysis in terms of artefact-free data duration or usable trials are not met). The problem of data attrition is compounded for naturalistic paradigms (such as free play), and also for dyadic protocols in which the computation of synchronicity measures requires both partners to concurrently contribute artefact-free data segments. As shown in Table 2, even when infants are watching engaging screen-based stimuli (such as cartoons or nursery rhymes), infant motion (including facial, limb and postural movement) is still present

between 60-70% of the time. In a naturalistic free-play paradigm, motion is present almost continuously. This high rate of motion contamination strongly motivates the development of robust algorithms for data cleaning, and a better understanding of the specific effects of different classes of motion on the neural signal (see section 3.2 for further discussion on these points).

Table 2. Prevalence of infant motion during four different experimental paradigms expressed as a percentage of the total experiment time.

Stimulus	Cartoon clips <i>N = 11</i>	Speech (nursery rhymes) <i>N = 24</i>	Speech (artificial language) <i>N = 14</i>	Tabletop free play <i>N = 5</i>
Average prevalence of infant movement (% of total stimulus presentation time)	59.2%	70.1%	71.0%	>95%

Note. The first three experiments were screen-based (i.e. infants watched either cartoon clips, natural speech or artificial speech respectively), whilst the last experiment involved free play across a tabletop with mothers. Infants were video-recorded during each paradigm and the recordings were manually coded for segments free of muscle movements of any kind. Here we report the average (across all infant participants) amount of time while performing the task and moving simultaneously. Authors' unpublished data.

Solution:

The usable fraction of infant neural data will depend on a range of factors, including the nature of the experimental paradigm (e.g. screen-based or interaction-based), the age of infant participants (older infants are capable of a wider range of motion), the sensitivity of the

recording device to motion (and effectiveness of motion artefact removal techniques), and the spectral and topographical properties of the neural effect of interest. In the case of an EEG dyadic interactive paradigm involving infants (e.g. Leong et al, 2017; Wass et al, 2018; Leong et al, 2019), we have previously noted a data attrition rate of ~50%. This suggests that studies may need to recruit and test twice the number of participants (dyads) than would be required based on standard power calculations. In the future, the advent of next-generation brain imaging systems such as Optically Pumped Magnetometer (OPM) MEG systems (Boto et al, 2018) which afford higher signal-to-noise ratios and greater robustness to motion artefact contamination, and better motion artefact cleaning methods (see Section 3.2) could permit high-fidelity and lower attrition neural data collection even in real-world naturalistic paradigms.

1.2 Varying arousal in younger infants

Problem:

Both longitudinal and dual EEG data can be affected by variations in infant arousal state. Infant behaviour can range from a highly agitated state (hyperarousal) to peaceful rest or even sleep in just a few minutes. Due to the relatively high ratio between sleeping and waking hours in neonates and infants as well as numerous EEG artefacts linked to neonate-typical motor activities during waking, such as sucking, crying, defecating and moving limbs (see Section 3.2), most of the EEG measurements of newborns are carried out during “quiet” non-rapid eye movement (NREM) sleep. However, while recordings during sleep improve the EEG signal-to-noise ratio, the obtained results may not be straightforward to interpret. In adults, even a mild decrease in arousal shifts the brain into a different mode of processing compared to the waking brain, affecting EEG event-related potentials (ERP), spectral power,

and connectivity measures (Ogilvie, 2001). Consequently, EEG results obtained during drowsiness or sleep cannot be directly interpreted in terms of brain functioning during waking hours. Early studies in infants suggested that neurobehavioural states of arousal (waking, NREM sleep and REM sleep) do not affect ERP characteristics in newborn infants (Hirasawa et al., 2002; Martynova et al., 2003), justifying infant testing during drowsiness and quiet sleep. However, subsequent research soon demonstrated such effects, including a significant increase of auditory ERP amplitude during quiet sleep compared to awake periods in 8 week old healthy infants (Friedrich et al., 2004; see Figure 2), and a change in the topography of mismatch responses as a function of alertness (Otte et al., 2013).

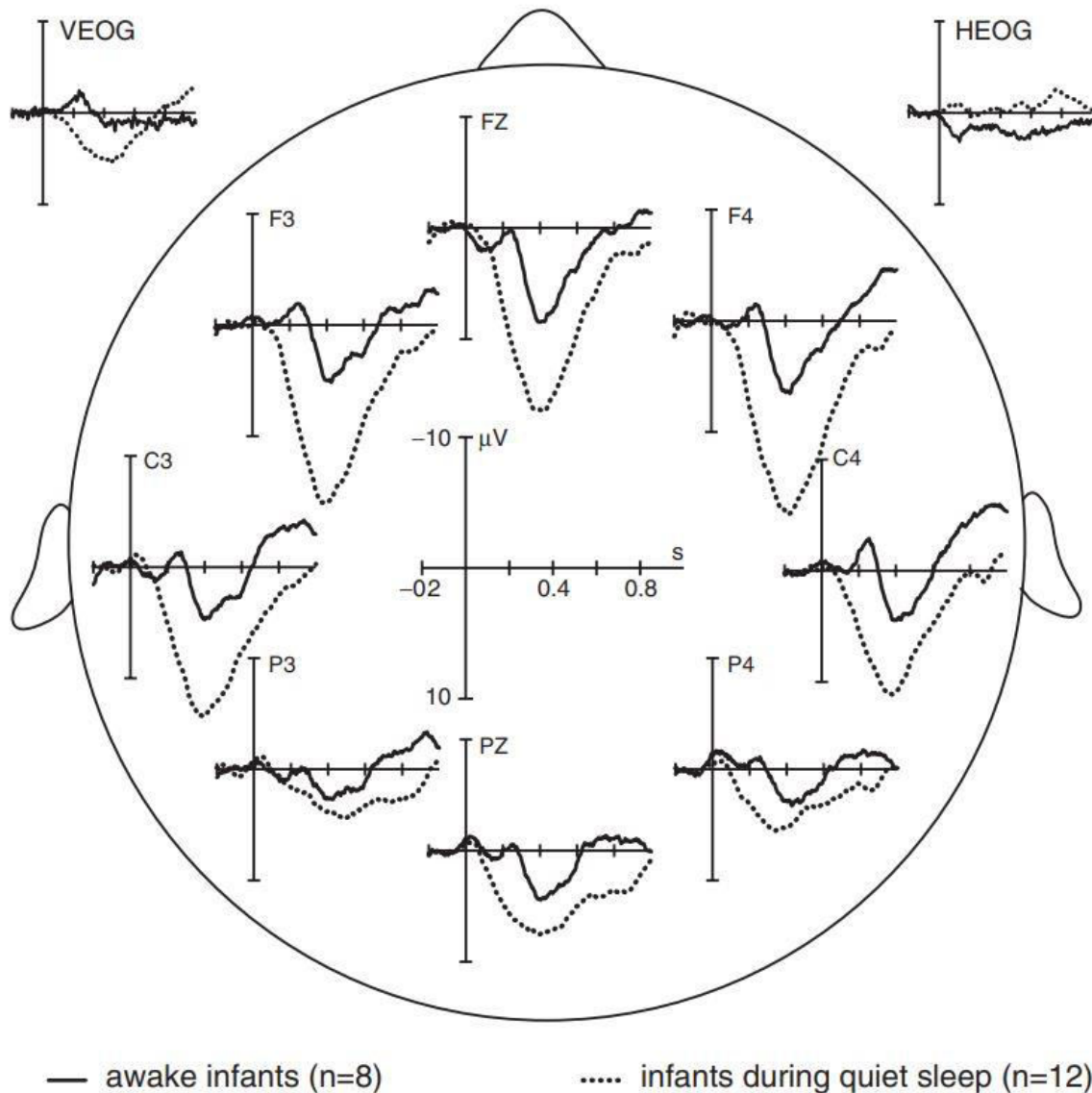


Figure 2 | Auditory difference waves (deviant minus standard) in 8 weeks-old infants show an increase of amplitude in quiet sleep compared to the awake state. Reproduced with permission from Friedrich et al. (2004).

Thus, infant EEG findings obtained during low levels of arousal should be interpreted cautiously when compared to alert datasets. The sleep confound is particularly problematic for longitudinal or cross-sectional EEG studies that compare younger sleeping infants with older awake infants. In such cases, an identified EEG change as a function of age may simply reflect EEG differences between different levels of arousal. EEG recordings carried out

during sleep for all age groups would not necessarily solve the problem, as the characteristics of sleep EEG change substantially during the first year of life, e.g. multifocal sharp waves are frequently observed during sleep in the neonates but not in older infants (Hagne, 1972). Similarly, dual EEG measurements can be confounded if one partner of a dyad is more agitated or drowsy than the other partner. Given that changing levels of arousal can shift background EEG activity towards faster or slower frequencies, this would likely reduce cross-brain synchronisation indices.

Solution:

A stable level of infant's arousal can be maintained by running a series of short 1-3 min experimental blocks with breaks lasting as long as needed in order to guide infants' arousal to the desired level. It is also critical that experimental scripts for stimuli presentation and data acquisition can be delivered in a contingent manner in response to changes in infants' behaviour and arousal state. Finally, data containing episodes of hyper- or hypo-arousal should be analyzed cautiously, either removing them or sorting into arousal bins and analyzing them separately.

When data acquisition during awake-only periods is not feasible in longitudinal or cross-sectional studies, such as when testing newborns, arousal level can be assessed in 20-30 sec epochs by manually scoring video recordings of infant's behaviour and/or scoring the level of arousal (awake, NREM sleep, REM sleep) from EOG and EEG recordings. Grigg-Damberger et al. (2007) summarise guidelines for the visual scoring of arousal and sleep in infants and children. There is also an increasing use of automated actigraphy-based assessment of sleep/wake patterns in paediatric research (Meltzer et al., 2012). The obtained arousal indices could be entered as covariates in statistical models testing experimental hypotheses.

2. Data acquisition

2.1 Determining the optimal location for reference and ground electrodes

Problem:

There are two major considerations to take into account when choosing a reference and the common ground (if applicable) electrode locations for infant EEG acquisition. The first consideration (more relevant to the common ground electrode if a given system uses such) is stability - choosing a location that is least vulnerable to motion-related displacement. The second consideration (more relevant to the reference electrode) is the topography of the effect that is being studied and comparability (in terms of the location and polarity of the predicted effect) to previous literature. Both of these considerations are also relevant to adult EEG, but adult participants are amenable to instruction (e.g. to stay still) and can more easily tolerate electrodes attached to different locations (e.g. nose, mastoid, and ear lobe). The issue is additionally complicated by developmental changes within the infant population.

Solution:

Regarding the common ground, the exact location on the participant is of relatively low consequence (Light et al., 2010), as long as the electrode is well attached and is not easily displaced by movement, e.g. when placed on the head or over a vertebra on the back of the neck. If an EEG cap is used, a recording channel is sometimes sacrificed for the common ground, which may not be ideal in a low-density system. In these cases, the back-of-the-neck vertebra is an acceptable alternative, as long as care is taken to attach it stably, avoiding its loosening during the recording session. For this, it should not be placed too high as that

would make it vulnerable to head movements, or too low, where it might pick up physiological movements such as breathing and heart-beat. The length of the wire should also be carefully considered, as it should give enough leeway for the infant to be able to lean forward without compromising the attachment of the actual electrode. Importantly, unlike with adults, the nose is a poor location of choice for infants (even neonates) given its high visibility and the potential for distraction.

The optimal location of the reference electrode has generated heated debate in the literature. Most amplifiers do not distinguish between recording and reference electrodes (Nunez & Srinivasan, 2006) - rather, they amplify the difference between pairs of scalp locations, and the resulting amplified potential is proportional to that difference. Since this manipulation is linear, in most systems the online reference can be recovered and the recording can be re-referenced offline. Additionally, as EEG is recorded as a difference of potentials, in an ideal world, the reference electrode would be symmetrically spaced from all other channels, and it should not record brain activity, so as to not have any brain activity subtracted out in the referencing process. Thus, for adults, the nose, earlobes, and mastoids (although they are on the head, the bone behind the temporal bone is thicker and provides a larger distance from the brain compared to the rest of the scalp sensors) are preferred sites (Op de Beeck & Nakatani, 2019). However, these locations are not easily tolerated by infants, and different locations are more tolerable for different ages, which immediately poses a problem in dual EEG designs, as placing reference at different locations for each partner of the dyad would affect EEG data differently. The mastoid location is well tolerated during recording, but placing reference electrodes at this location often causes distress to infants during removal because adhesive stickers that are used to attach the electrodes securely can easily pick up longer hairs.

Therefore, careful consideration should be given to the safe and speedy removal of electrodes from mastoid locations for infants.

Perhaps the most frequently applied solution is to use Cz as a reference channel (Heohl & Wahl, 2012). Its location is symmetrical to all other channels, it is a fairly protected location from motion artefacts, and it allows for shorter setting up and removal time, especially if a cap is used. This solution has three potential pitfalls, however. First, it sacrifices a recording channel, which is not desirable in low-density systems. Second, if the infant has thick hair growth, getting a good connection and low impedance might be difficult. Finally, the central location is maximal for certain forms of brain activity, and this activity would be subtracted in the referencing process. Evidently, there is no perfect solution when it comes to the location of the reference channel in infant EEG. The decision should be informed by considerations such as the setup (high or low density); the age of the participants; the experimental question and design (is it a dual EEG study); and finally and most importantly, the published literature - changing the reference location can change the direction of the dipole that is being measured, and the topography of effects.

2.2 Control of impedance when testing with the high-density sponge electrode nets

Problem:

Of the different equipment setups available to researchers who wish to record infant EEG, one important distinction is between saline- and gel-based systems. For gel-based systems (such as those from Biosemi, Neuroelectronics, and Biopac), each electrode must have electrolyte gel manually inserted after application of the net or cap before recording can start. For saline-based systems (such as the EGI/Phillips system) each electrode contains a small

sponge, and the entire net is soaked in electrolyte before it is applied to the infant's head. The advantage of these systems is that, because each electrode does not need a manual injection of electrode gel, high-density recordings (e.g. with 128 electrodes and more) can be taken from infants, who otherwise would not tolerate a long process of applying a high-density net prior to recording.

One disadvantage of a saline-based system is that the connectivity between electrodes and scalp relies on the presence of the saline solution. As the recording progresses, and the saline solution evaporates, impedances can increase over time, systematically decreasing data quality towards the end of the recording session.

Solution:

First, recording sessions using saline-based nets should be kept brief (e.g. no more than ~20 minutes per continuous session). If longer recordings are necessary, testing can be interrupted and the saline solution reapplied, electrode by electrode, using a pipette, and without removing the net. Second, if comparisons are planned between conditions, ensure that the order in which conditions are presented is counterbalanced between participants, to avoid the possibility that systematically higher impedances later in a testing session may influence results. Finally, if planned analyses will focus on a subset of electrodes, ensure that these electrodes are manually checked, by viewing the impedances both before and during testing.

2.3 Poor fit of standard EEG caps to the varying shapes of infants' heads

Problem:

Poor fit of standard EEG caps is a common problem for measurements with young infants. This arises in part from the fact that the fontanelles (“soft gaps” between the cranial bones made of connective rather than bone tissue) are still malleable during early infancy and therefore the shape of the infant’s skull (therefore, head), can change dramatically, especially in the first week of life (Baum & Searls, 1971). Furthermore, changes in infants’ head shape can be non-linear in the first weeks of life of pre-term infants (Souza et al., 1976; see Figure 3). This can be problematic in longitudinal study designs when an infant is tested at birth and then again later in life, as the exact location where the electrodes lie using a standard cap can vary slightly between the ages. This may or may not pose a problem depending on how sensitive the research question is to the topography of the effect. The fontanelles normally close up slowly starting at about 3 months, with the process continuing to 18 months and beyond (Duc & Largo, 1986; Hansman, 1966), but these gradual changes are small and are only likely to be problematic with very low density setups, i.e. fewer than 8 recording channels.

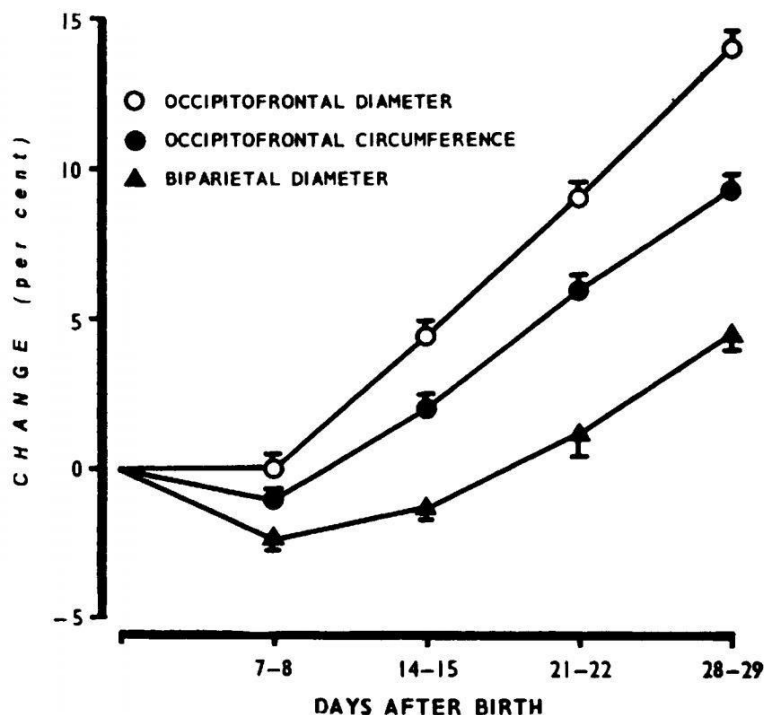


Figure 3 | Percentage changes (y axis) of the head shape during the first month in preterm infants born by vaginal delivery at 26-35 weeks' gestational age. Reproduced with permission from Souza et al. (1976).

Solution:

In a clinical setting, where precision in topography and electrode placement are essential, the most frequent use of EEG involves only one or two recording channels (plus a reference/common ground). These electrodes are individually glued using a conductive paste to a specific location on the scalp irrespective of the shape of the infant's head (e.g. for monitoring seizures).

In neonatal and early infancy research, a low-density (typically between 3-32 channels) setup is more practical and likely to be used than a high-density setup (32-128 channels) (DeBoer, Scott & Nelson, 2007). With the lower number of channels to be attached, they can be applied individually, though time and care should be taken to correctly identify each channel's location, which depending on the setting and the state/mood of the infant, can be more or less practical. The manual application of channels may be more prone to human error than the application of a cap. However, given the high variability in head-shapes of infants < 3 months old, if the experimenters are well trained and experienced, and the exact rather than approximate location of channels is essential, individual application of channels may be prescribed. In high density systems, however, this would simply not be practical. In infants older than 6 months, where generally higher-density (32+ channels) systems are preferred, and where there is little indication of greater head-shape variability than in adults, using a stretchy cap is appropriate standard practice.

2.4 Ensuring a high degree of temporal synchronisation between amplifiers and recording devices – dual EEG acquisition

Problem:

Data analyses for dual EEG data require that the different recording devices used have a very high degree of temporal synchronisation. When measuring brain activity at 10Hz, for example, for which one complete oscillation lasts 100ms, a variable recording error of even tens of milliseconds can be detrimental in estimating the phase of a signal. Similar considerations apply to time synchronisation between video recordings and EEG data. Our experience suggests that certain EEG recording systems, and many cameras, can periodically skip short segments of data during recording, particularly in cases when the data sampling rate is high and the memory/hard disk write speed of the system is limited.

Solution:

First, ensure that, wherever possible, one amp rather than two separate amps is used for the EEG data acquisition. This eliminates problems of ensuring accurate synchronisation between the two (or more) EEG data-streams. If this solution is not possible, such as when synchronising the EEG data with external video cameras, then a regular TTL pulse may be sent (from a separate computer) which inserts a periodic event into the EEG recording(s), and simultaneously sends a signal (e.g. by lighting an LED) which is visible to the video cameras. Synchronisation of the EEG to the video can be ensured by hand-coding the frames in which the LED light is first visible. Rather than synchronising once, this same process should be repeated throughout the recording, in case of short periods of missing data (e.g. dropped frames) mid-way through recording.

3. Analysis of longitudinal and social neuroscience datasets – single participant data

3.1 Developmental changes in head anatomy affect EEG source localisation

Problem:

Major advancements are being made in computational methods for modelling neurodevelopment changes of sources of EEG signals during infancy (e.g. Reynolds & Richards, 2009; Xie & Richards, 2017). However, this is challenging because significant changes in head anatomy occur during the first years of life, which affect skull conductivity and the propagation of electrical signals that are measured at the scalp (Odabae et al., 2014). For several months following birth, the neonatal cranial bones are not fully fused. Rather, the skull bones are separated by fontanels (soft membranous sutures) that accommodate the stretching and deformation which occurs during birth and subsequent growth of the brain (Weickenmeier et al., 2017). Fontanels are initially relatively small during delivery but can expand up to 3-4 cm in width in the first few months of life (Popich & Smith, 1972; Pedroso et al., 2008). In healthy children, the anterior fontanel (the most prominent diamond-shaped membrane-filled space located between the two frontal and two parietal skull bones) is generally the last to close (due to ossification of the bones of the skull). The anterior fontanel persists until around 18 months after birth and sometimes for several years (Hansman, 1966). Fontanels influence the volume currents that accompany primary currents generated by active neuronal populations. Modelling data suggest that the influence of fontanels is akin to that of holes (Benar & Gotman, 2002) and zones of higher conductivity (Ollikainen et al, 1999) in the adult skull. Specifically, Roche-Labarbe et al (2008) reported that in their source localisation analyses, when the infant fontanel was modelled as a thinner resistive skull zone, this caused a modest dipole shift towards the fontanel region.

As the thickness of the skull increases from 1 to 2 mm at birth to several millimetres during early adulthood, its conductivity concomitantly decreases (Gibson et al, 2000; Hansman, 1966). In adults, there is high individual variability in skull conductivity (Goncalves et al, 2003), which suggests that infant skull conductivity may vary considerably across participants as well. Infant source localisation modelling studies have shown that changes in skull layer conductivity tend to affect the magnitude and position of dipoles to a larger extent than their orientation (Roche-Labarbe et al, 2008; Lew et al, 2013).

Solution:

Currently, most available source localisation techniques for infants do not compensate for factors such as skull thickness, density, and fontanel closure. Although high-density EEG arrays may offer promising potential for estimating sources of neural activity in infants, it is recommended that attempts at source localisation are combined with, and constrained by, strong theoretical and anatomical predictions. In particular, attempts to develop developmentally-appropriate head models (e.g. Lew et al, 2013) may improve the feasibility of source localisation for infants in the future. Azizollahi et al. (2018) have recently developed promising head models for the forward solution in EEG source localisation that take into account variable conductivity of fontanels; follow up research is needed to test the proposed solutions at different stages of fontanel closure.

3.2 Prevalence and diversity of infant myogenic artefacts

Problem:

For paediatric populations, motion is almost unavoidable during experimental recordings. In EEG, myogenic artefacts are more problematic for the infant than adult measurements since involuntary physiological activity such as heartbeat and blinks are less stereotypical and therefore more difficult to identify and remove using the de-noising procedures normally applicable to adult EEG (Fujioka, Mourad, He, & Trainor, 2011). A further complication arises from infants' tendency to move abruptly and frequently, which creates temporary displacement of channels on the scalp (Bell, & Cuevas, 2012; Fujioka et al., 2011; Hoehl & Wahl, 2012).

As demonstrated in Table 2 (Section 1.1), the prevalence of movement during the time on task can be up to >95% in naturalistic interactive paradigms. Muscular activities can be grouped in the head-neck-face category and the bodily-limb movement category. We generally find that the former is more problematic in terms of prevalence, and in terms of broadband artefactual effects on the EEG signal.

Two representatives of the first category worth mentioning are eye and jaw movements. They occur from birth (i.e. blinking, sucking reflex, yawning, rooting, crying) and continue to occur frequently in experimental conditions by 18 months and beyond. The proximity of muscular generators of these motions to the EEG electrodes and their repetitive nature makes them particularly problematic for the EEG researcher. While the effects of eye-blinks on EEG are somewhat better studied in the literature and can be detected using a dedicated channel, or eye-tracker (Plöchl, Ossandón, J., & König, 2012), jaw motions are more challenging to both detect and correct, and are known to reduce the signal-to-noise ratio of neural signals that relate to cognition (Brooker & Donald, 1980).

In the body-limb movement category, we find that upper limb movements (arms and hands) create the most artefactual spectral signature in infant EEG during naturalistic play in the age group around 12 months (see Figure 4). A problematic feature of these sets of motion artefacts is that they do not necessarily produce obviously artefactual activity in raw EEG - i.e. the amplitude and temporal features of the electrode traces may not change much compared to when the infant (although this is also true for adults) is apparently still. Therefore, contamination from these forms of muscular activity may remain undetectable by automated pre-processing procedures, and indeed during manual artefact rejection (see Figure 5).

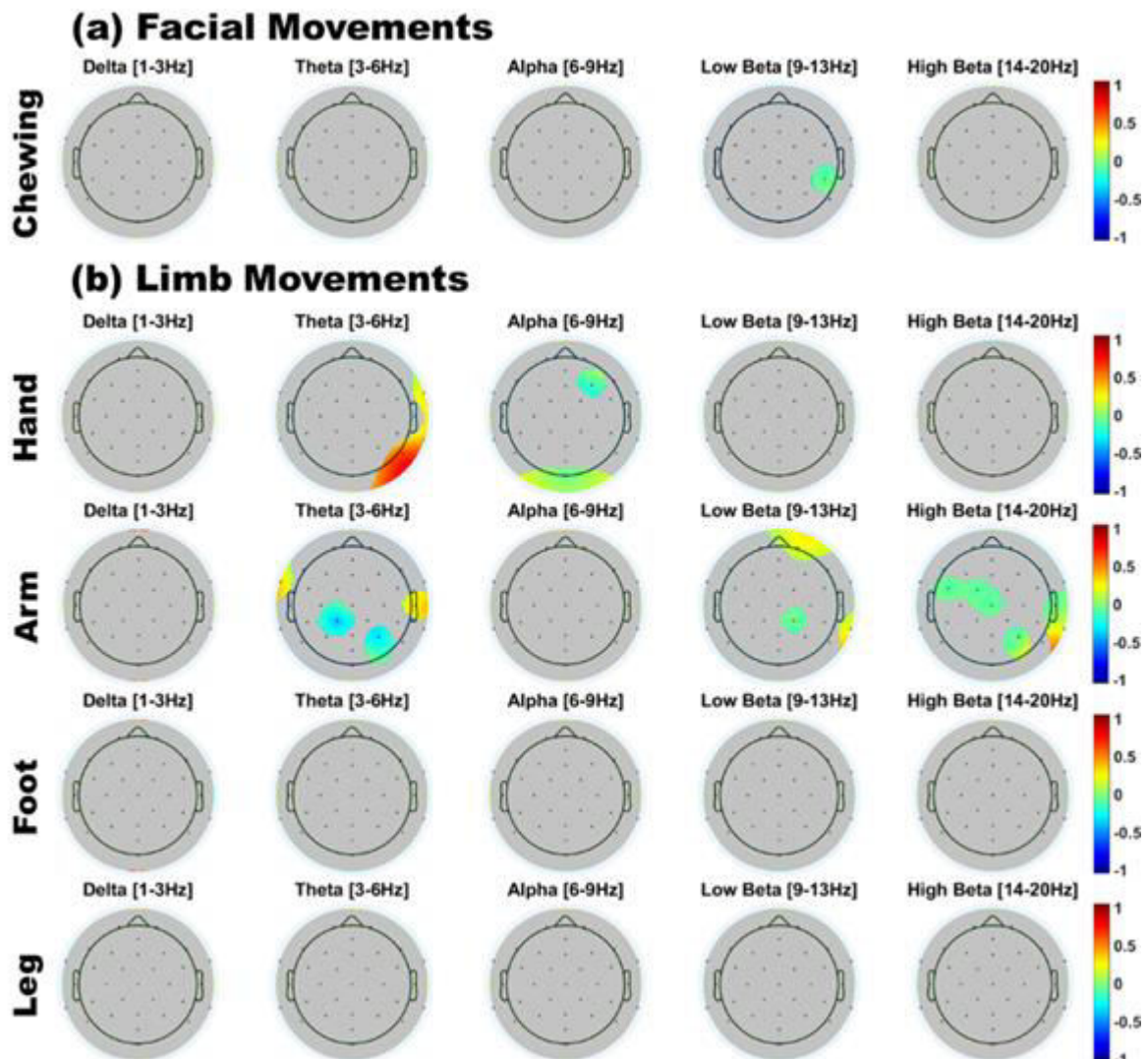


Figure 4 | Scalp topography of significant differences (Cohen's D-values) between a) facial movements or b) limb movements, and resting state. In each subplot, only scalp regions where significant power differences were observed (Benjamin-Hochberg FDR corrected at $p < .05$) are shown in colour. Grey = nonsignificant result; red = higher power during movement; blue = higher power during resting state. Fourteen infants participated in the study. There were 6 boys and 8 girls in the group, with an average age of 338.85 days (SD = 59.59). Two infants produced insufficient resting state data due to fussiness, and so were excluded from the analysis. The remaining 12 infants comprised 5 boys and 7 girls, with an average age of 325.5 days (SD = 51.77). Facial movements are represented by 6 infants, hand and arm movements – by 8 infants, and foot and leg movements – by 10 infants. The study was approved by the Cambridge Psychology Research Ethics Committee, and parents provided written informed consent on behalf of their children. Authors' data (Georgieva et al., 2018).

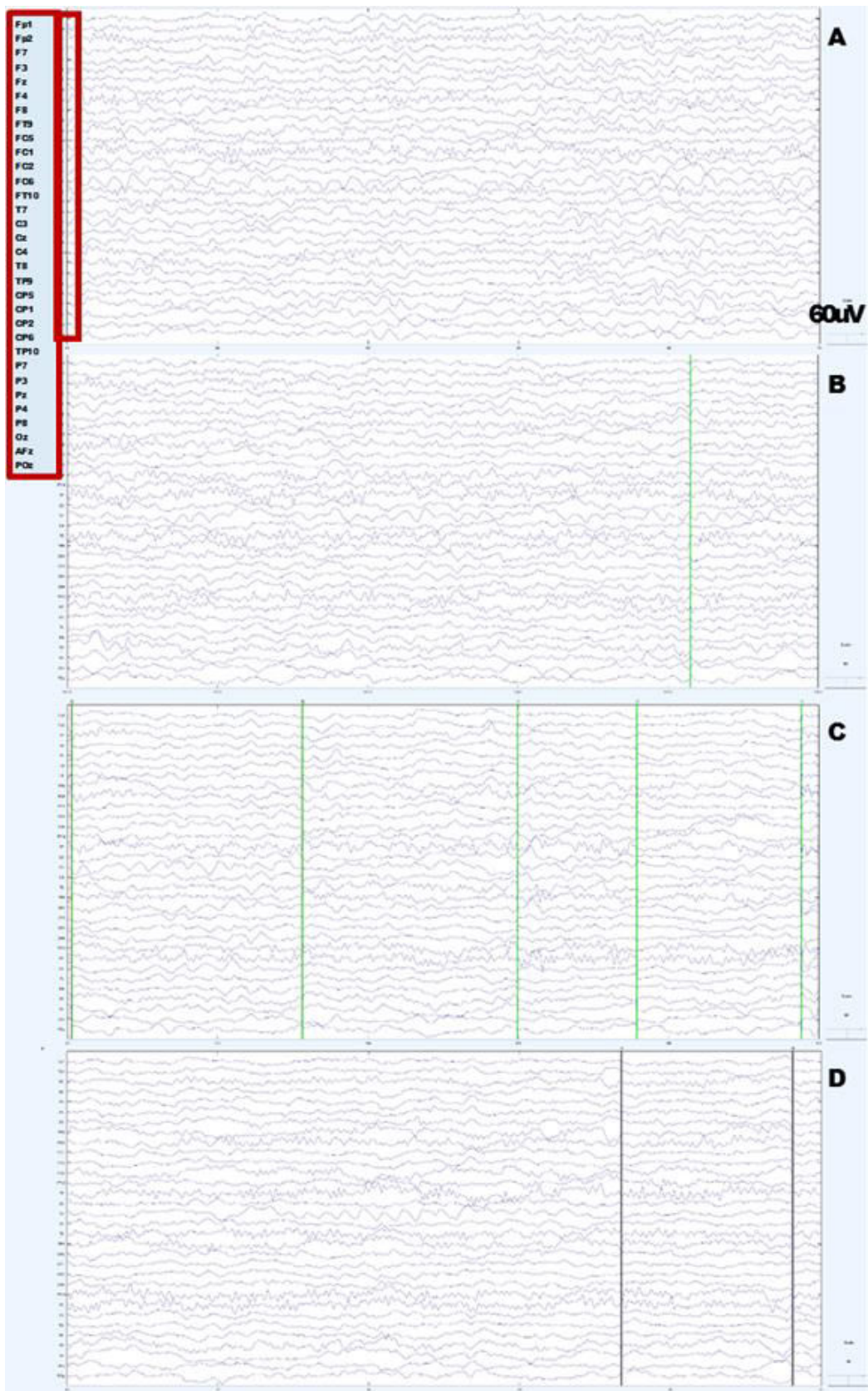


Figure 5 | Infant raw EEG (filtered and average re-referenced) segment from continuous recording. All figures show 5 seconds of continuous recording during a) infant resting state, b) infant hand movements, c) infant arm movements, and d) infant talking/babbling. Note

that timepoints where movements were made (marked as vertical lines) do not produce easily identifiable changes in the EEG signal. Authors' unpublished data.

In cases where the electrical trace displacement is not as stereotypical as with blinking, one could try to isolate signal distortion by comparing its frequency spectra with that of resting state activity. However, when interpreting muscle generated artefactual activity in EEG, using infant “resting state” activity as a gold standard is questionable in its own right. It is difficult to ascertain with confidence that an infant’s EEG recording is movement free; even when there is no apparent motion. For example, there might be gentle continuous sucking that occurs at ~2-4Hz, which would increase delta or theta power in temporal electrode regions. Such activity would be virtually undetectable without additional EMG electrodes over the jaw area - and these are nearly never used in infant EEG research.

Solution:

As there is no efficacious method for eliminating infant motion during EEG experiments, it is essential to understand (1) the characteristic effects on typical infant motions on the EEG signal, (2) the motions most likely to be exhibited at the age of interest, (3) the motions most likely to be present in the paradigm, and (4) the effect of motion on different EEG metrics (e.g. connectivity or spectral measures).

To address the problem of movement-related artefacts in adult EEG data, researchers traditionally exclude artefact-contaminated segments by employing strict rejection thresholds. However, when applied to infants, the artefact exclusion approach may entail considerable data loss (Fujioka et al., 2011), and there is an increasing interest in correction procedures that permit the accurate identification and removal of artefacts from infant EEG without

significantly compromising the integrity of underlying neural activity (Gwin et al., 2010; He, Hotson, & Trainor, 2007; Fujioka et al, 2011). Yet, while various methods can ameliorate the problem of data loss due to artefact rejection, and they have been used in a number of research studies as a cleaning strategy for data pre-processing (Agyei, Weel, & Meer, 2016; Corrigan & Trainor, 2014; Folland, Butler, Payne, & Trainor, 2014; Slugocki & Trainor, 2014; Trainor, Lee, & Bosnyak, 2014), these strategies still rely on the successful classification of portions of the EEG signal as artefactual, which is itself a non-trivial task.

In order to assess whether the artefactual activity is likely to inflate or reduce the effect that researchers are testing, systematic control analyses should be carried out on how specific movement artefacts distort infant EEG data. We have successfully employed this strategy with jaw movements artefacts in a study where an adult sang nursery rhymes for the infants and her EEG was recorded during speech production (Leong et al, 2017). We quantified the spectral and topographical signatures of the production of nursery rhymes and compared those to a silent resting state in order to quantify the effect of speech production artefacts on our measures of interest (connectivity patterns). Furthermore, in more recent work, we attempted to catalogue a detailed descriptive library of the most prominent motions produced by infants, and their respective artefactual influence on the EEG recordings (Georgieva et al., 2018). Accordingly, we investigated their topological and spectral features as compared to resting state EEG in the major frequency bands of interest (0.05 - 20 Hz). While this library is by no means exhaustive, we hope that it can be further developed as a tool to identify which dimensions of EEG signal are the most vulnerable to biases introduced by artefactual muscle activity.

3.3 A lack of stereotypical topography of infants' eye movements needed for the ICA-based data cleaning

Problem:

Given that infants cannot be instructed to avoid blinking during critical periods of stimuli presentation, infant task-related EEG data can become contaminated with a much higher proportion of eye-movement related noise than adult EEG. This poses data comparability issues for dual EEG studies that require a comparable signal-to-noise ratio for the accurate estimation of cross-brain synchronisation. Stereotypical EEG artefacts, such as those caused by blinks and saccades, can be efficiently identified and removed from the signal using Independent Component Analysis (ICA), which separates a multivariate signal into additive subcomponents that are linearly mixed in several EEG channels (Makeig et al., 1996; Vigario, 1997). In adult EEG research, ICA-based decomposition of signal is perhaps the most commonly used method of data cleaning, and there is also an increasing use of ICA in infant research. However, while ICA is very efficient in identifying eye-movement related artefacts in the adult populations, it often fails to identify components driven by blinks and saccades in infants (see Figure 6). In some other cases of ICA of infant EEG, eye movement noise spreads across a large number of components (Fujioka et al., 2011), as it is not as temporally and spatially confined as in adults (Bell & Wolfe, 2008). Also, given that infant background EEG activity is relatively slow in frequency but high in amplitude (Bell & Wolfe, 2008), its spectral and spatial patterns can overlap with those caused by eye movements, further obstructing ICA-based decomposition of the signal.

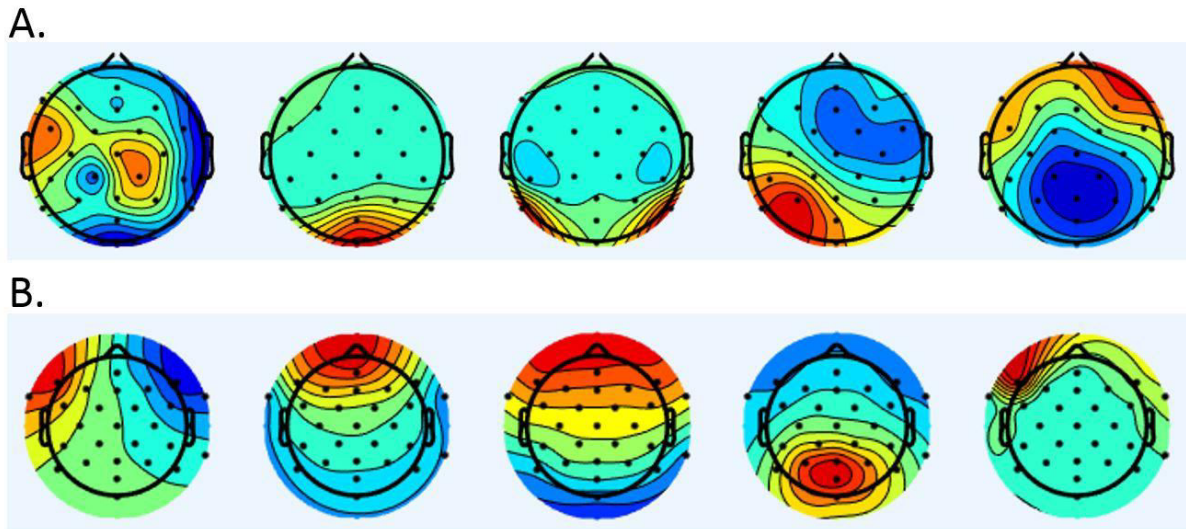


Figure 6 | Topographic representation of the first five independent components, ordered from the largest to smallest by the magnitude of EEG signal explained, calculated from a) infant (9.5 months) and b) mother (32 y.o.) datasets of the same dyad. The first two components in the mother’s dataset represent eye movements (saccades and blinks), whereas eye movement-related components are missing from the ICA decomposition of the infant’s dataset. Authors' unpublished data.

Solution:

While most of the EEG studies using ICA for data cleaning rely on the Runica (Makeig et al., 1996) or FastICA (Hyvärinen & Oja, 1997) algorithms, there is a large body of signal processing literature investigating other ICA solutions, and some of them could be more efficient for the processing of infant data. For instance, Miljkovic et al. (2010) compared the efficiency of RobustICA, SOBI, JADE, and BSS-CCA algorithms to extract ECG artefacts from neonatal EEG signal, and found their sensitivity to vary from 0.73 to 0.85 with SOBI and BSS-CCA methods being more efficient than RobustICA and BSS-CCA. A similar systematic research is needed to compare the efficiency of different ICA methods for cleaning both infant and adult EEG datasets. Finally, when ICA-based algorithms fail, other artefact cleaning methods should be applied, such as Artefact Blocking algorithm, which can

efficiently process both infant and adult datasets (Fujioka et al., 2011; Mourad et al., 2007). It multiplies the observed data matrix by a blocking matrix that eliminates high amplitude artefacts, such as blinks, saccades or head movements, while linearly transforming the other EEG sources with minimal distortion.

3.4 Relatively high inter-individual variability of EEG responses in younger cohorts

Problem:

In EEG studies with infants, two inherent problems are that shorter periods of data recording must be used, and that data quality during these recordings is more variable, due to the increased prevalence of artefacts. In addition, one further problem is that intra-individual variability in neural responses is inherently higher in infants (De Haan, 2013). This higher variability arises from a number of sources, including variable arousal states (see section 1.2), movement artefacts and non-compliance (see section 3.2), and rapid developmental maturation of brain structures (see sections 2.3 & 3.1). This high variability means that a larger number of samples must be taken to obtain an accurate average value. And, finally, due to differential rates of development and brain maturation, inter-individual variability in responses is also higher in younger cohorts. Taken together, these problems pose serious signal-to-noise challenges in infant EEG studies.

Solution:

Practical steps that can be taken to address these problems are: i) aiming for a smaller number of conditions, allowing more (and therefore more accurate) data to be collected for each condition; ii) splitting data collection, where possible, across a number of different shorter sessions, or interspersing an experimental condition with engaging TV clips, in order to

maximise data retention. Analytical steps that can be taken include inserted the amount of data collected in as a covariate into statistical analyses.

4. Analysis of dual EEG data

4.1 Developmental shifts in canonical EEG rhythms

Problem:

During the early years, rapid developmental changes in brain activity occur (Haartsen et al., 2016), and canonical EEG rhythms change over development (Anderson & Perone, 2018; Hagne, 1968). For example, Stroganova and colleagues (1999) found that alpha activity is slower during infancy than during early childhood and later in life (see Figure 7A), although it functionally and topographically corresponds to posterior alpha in adults, and that theta activity is also lower (Orekhova et al., 2006). Marshall and colleagues (2002) found, similarly, changes in relative 6-9 Hz power from 5 months to 4 years of age, with peak frequency increasing from 4 Hz (over parietal and occipital sites) and 7 Hz (over central and frontal sites) at 5 months to about 8 Hz by 2 years and 9 Hz by 4 years (widespread over the scalp) (see Figure 7B). They also observed stable individual differences in relative alpha power from 10 months to 4 years (but not 5 months to 10 months).

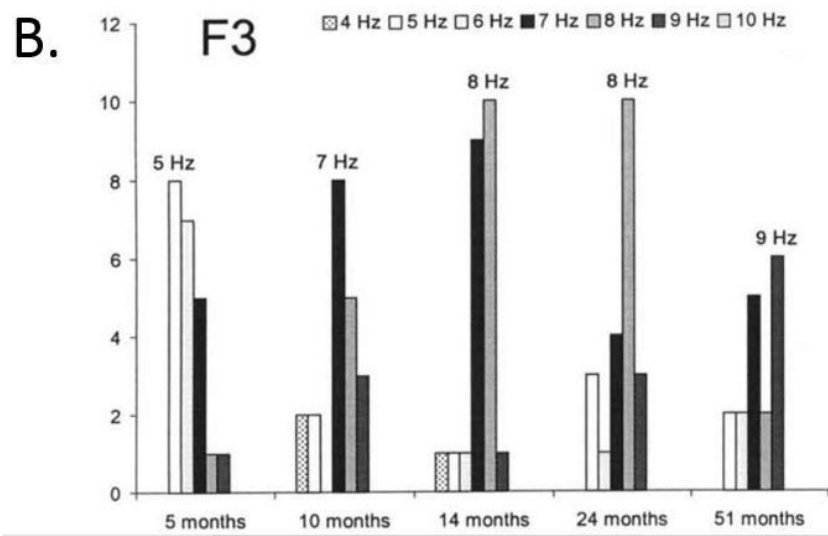
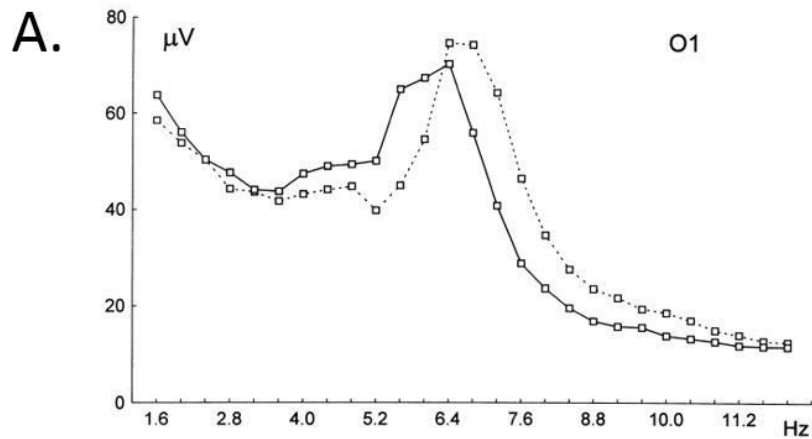


Figure 7 | (A) A developmental shift in the alpha frequency range observed from the age of 8 months (solid line) to the age of 11 (dashed line). Reproduced with permission from Stroganova et al. (1999). **(B)** The number of participants from five age groups depicted as vertical bars, whose peak frequency fell at a specific 1 Hz bin within the 4–9 Hz range, recorded from the left frontal electrode F3. A peak frequency shift can be observed from 5 Hz at the age of 5 months to 9 Hz at the age of 51 months. Reproduced with permission from Marshall et al. (2002).

Such developmental shifts in EEG rhythms present problems for dual infant-adult EEG analyses, as most of the analytical algorithms used to study inter-brain connectivity, such as phase locking value (PLV; Lachaux et al., 1999) or weighted phase lag index (wPLI; Vinck et al., 2011), estimate phase coupling of identical frequencies extracted from two time series.

Given that peak frequencies of oscillatory rhythms are slower in infants compared to adults, such analyses are likely to assess phase coupling of different frequency bands between dyad partners. For instance, a 6 Hz coupling between an 8 month old infant and her mother would estimate synchronisation between alpha oscillations in infant and theta oscillations in mother. Given that different frequency ranges may have different functional roles, such coupling may be suboptimal for the identification of functionally significant inter-brain synchronisation.

Solution:

As an alternative to the single-frequency coupling, inter-brain connectivity can be assessed using cross-frequency coupling techniques (Canolty & Knight, 2010; Cohen, 2014). In the phase-amplitude coupling analysis, an oscillation phase in one time series is correlated with an oscillation amplitude in another time series. Importantly, the frequency of oscillations does not need to match between the two time series, enabling coupling between different frequency bins, e.g. 6 Hz alpha phase in infant and 10 Hz alpha amplitude in mother. Likewise, the power of different frequency bins can be correlated between two time series (infant and mother) in the amplitude-amplitude coupling. Alternatively to the cross-frequency coupling methods that rely on the analysis of narrow frequency bins, inter-brain connectivity could be assessed using connectivity indices that are not limited to a single narrow frequency bin. For instance, weighted symbolic mutual information (wSMI; King et al., 2013) estimates the probability of co-occurrence of symbolic signal patterns in two times series at a rather low frequency specificity. While symbols formed by three elements separated by 32 ms have peak sensitivity at 8 Hz, they represent ~2-11 Hz range of frequencies, which would arguably cover alpha peaks for both infants and adults (King et al., 2013).

4.2 Differentiating inter-personal synchrony from spurious synchrony due to common intrinsic properties of the signal

Problem:

Dual neuroimaging studies usually examine synchrony in one of two modes. The first is concurrent synchrony ('when X is high, Y is high'), which can be assessed either by measuring covariance in the amplitude or power of the signal or by measuring the phase of the signal (e.g. phase-locking value; Lachaux et al., 1999). The second is sequential synchrony ('changes in X forward-predict changes in Y') (often known as Granger-causality; Granger, 1969) which can again be assessed based on the amplitude, power, or phase of the signal (e.g. Partial Directed Coherence; Baccalá & Sameshima, 2001).

One challenge to researchers measuring synchrony is that neural activity itself has certain common properties. For example, two adults, each with a dominant alpha rhythm of 10Hz, might be expected to show fairly consistent phase relationships between their alpha rhythms even if there was no communication between them. Similar considerations apply when considering variability in amplitude and power of the signal, where common oscillatory activity may be attributable to other sources (such as heart beat; Schmueli et al., 2007).

Solution:

Because of this inherent problem, most hyperscanning studies do not simply measure phase coupling between individuals, but rather compare the degree of coupling between different experimental conditions. Ideally, the two conditions should be held as constant as possible except for the particular parameter (e.g. the presence or absence of social interaction) which is of interest.

Second, from an analytical perspective, researchers (Burgess, 2013) have argued that, perhaps particularly for EEG, where analyses tend to focus on the oscillatory components, the simple presence of consistent phase differences between two EEG signals is perhaps too inclusive a definition of synchrony, and that Granger-causal measures such as Partial Directed Coherence may give fewer false-positive results (Burgess, 2013).

One recommended analytical solution is to perform a bootstrapping analysis in which corresponding epochs from the dyad are temporally translocated ('shuffled') with respect to each other, and synchrony is re-computed multiple times, in order to estimate the likelihood that the synchrony observed in the paired datasets would have been observed by chance. In doing this, care should be taken to ensure that sufficiently large epochs are used to ensure that lower frequency information is not distorted (which can be challenging in infant datasets, as any longer segments of usable data can be rare).

4.3 Differentiating inter-personal synchrony from spurious synchrony due to shared external perturbation

Problem:

The second problem facing researchers measuring inter-personal synchrony is that neural activity is also influenced by common external perturbation due to the environment, e.g. sounds or visual stimuli in the lab. For example, neural activity synchronises to temporal structures in speech (Giraud & Poeppel, 2012), and hence any words uttered during EEG recording could boost spurious synchrony by triggering similar neural responses in both partners of the dyad. Differentiating neural inter-personal synchrony from synchrony

attributable to shared external perturbation can be non-trivial (Fairhurst & Dumas, 2019). This is particularly true for naturalistic paradigms where factors such as the acoustic input cannot be completely controlled for across participants.

Similar considerations apply to analyses examining event-related change. Burgess showed that measures such as Phase-Locking value are sensitive to changes in the variance of the marginal distributions from the expected phase. Therefore, phase resetting time-locked to a stimulus can create the appearance of increases in phase synchrony, even when this is not the case (Burgess, 2013).

Solution:

A bootstrapping analysis in which chunks are translocated - either within a dyad over time or between dyads – does not control for this confound in cases where the environmental input is inconsistent across time, or across dyads. In cases where external stimulation cannot be adequately controlled for, the solution to this problem is non-trivial (Fairhurst & Dumas, 2019). As with section 4.2 above, an optimal experimental design will not just measure synchrony per se, but examine how synchrony changes relative to different experimental conditions – e.g. in the presence or absence of social interaction. Ideally, the level of external stimulation should be constant between these conditions. In addition, separate analyses can be conducted to identify the influence of external perturbation in each partner individually (e.g. speech-brain synchrony, in the case of an experiment where one or both of the partners is talking). And then, separately, an analysis can be conducted to show that speech-brain synchrony *does not* differ between conditions for either individual 1 or individual 2, but that brain-to-brain synchrony *does*.

5. Test-retest reliability of EEG measurement in infants

Problem:

The stability and reliability of a testing instrument over time is one of the critical requirements for reproducible experimental research. In response to the replication crisis in cognitive sciences, cognitive neuroscientists have begun to show a growing interest in the test-retest reliability of their paradigms and brain measures, and several recent studies have assessed the test-retest reliability of EEG measurements in infants.

Munsters et al. (2019) reported the test-retest reliability of ERPs evoked by neutral, fearful and happy faces presented in low or high-frequency conditions to 9–10 month-old infants (N=31), who were tested twice within two weeks. Substantial reliability was observed for the amplitude of the earliest N290 component (Intraclass correlation (ICC)=0.76), and moderate reliability was observed for the amplitude of components P400 (ICC=0.58) and Nc (ICC=0.57) (see Figure 8A). These findings suggest that the reliability of the early sensory potentials might be generally higher than the reliability of later potentials that reflect more complex and variable cognitive processes.

van der Velde et al. (2019) estimated the test-retest reliability of EEG connectivity and network characteristics in 10 month-old infants (N=60), who were tested twice one week apart while watching videos on a computer screen. The reliability of the global phase lag index (PLI) was excellent for theta and alpha frequency bands (3-12 Hz; intraclass correlation coefficient (ICC)=[0.84-0.91]), and mediocre to good for delta, beta and gamma bands (ICC=[0.60-0.72]). Similarly, the test-retest reliability for the graph measures (Average clustering coefficient and Characteristic path length) calculated from the PLI matrices was

excellent for theta and alpha frequency bands (ICC=[0.84-0.91]), and mediocre to good for delta, beta and gamma bands (ICC=[0.53-0.73]). However, considerably lower reliability was observed for the Small-worldness Index for all frequency bands (ICC=[0.13-0.67]). Thus, EEG connectivity and network characteristics can be reliably assessed in 10 month-old infants (see Figure 8B), although reliability is strongly dependent on the frequency band and specific graph index.

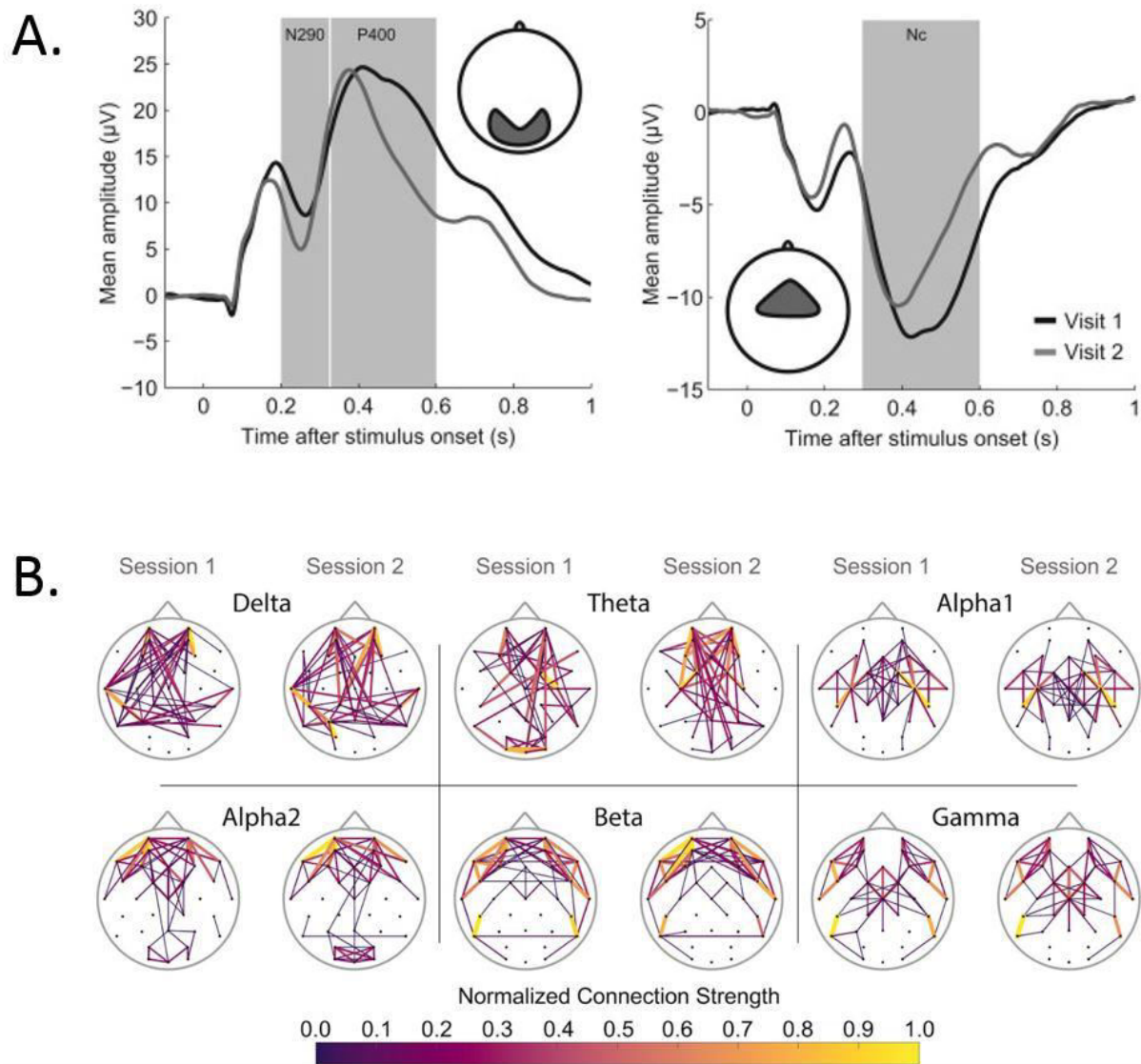


Figure 8 | (A) Areal average of ERP to face stimuli in 9-10 month-old infants show a larger amplitude decrease for the later components during the second session. Reproduced with permission from Munsters et al. (2019). (B) Averaged PLI connectivity maps with the 12% strongest connections in 10 month-old infants indicate great similarity between two sessions. Reproduced from van der Velde et al. (2019), following the Creative Commons Attribution License (CC BY).

Overall, these studies demonstrate substantial reliability of specific EEG measures in both active and passive tasks. However, given that ICC is not equally high across all ERP components, frequency bands, and graph indices, the reliability of any new or untested EEG measure cannot be inferred from the reliability of related measures. Furthermore, while

reliability was relatively high for the averaged ERP data (Munsters et al., 2019), specific experimental manipulations showed low reliability. In particular, even though there were significant ERP effects of face emotion and an emotion x spatial frequency interaction, these experimental manipulations yielded very low reliability (ICC=[-0.33-0.15]), suggesting that reliability of psychological manipulations may be much lower than the reliability of purely neurophysiological measurements.

Solution:

Ideally, the introduction of a new experimental manipulation and/or a new EEG analysis method should be accompanied by the estimation of its test-retest reliability. This is particularly important for the assessment of complex statistical interactions and new areas of research, such as dual EEG interactions between infants and adults.

Concluding note

Most of the EEG acquisition and analysis techniques have been developed to study adults, and hence their adaptation for infant research should always be carefully scrutinised. While longitudinal and social neuroscience studies using infant EEG become increasingly sophisticated (e.g. using dual EEG paradigms), new unique methodological challenges will continue to be identified. This requires infant EEG labs to invest time and resources into building technical and signal processing expertise that is not always available in the research centres dedicated to studying developmental psychology. With human cognitive neuroscience becoming increasingly specialised and technically and computationally challenging, a close collaboration between infant EEG and methods-oriented labs is desirable if not necessary.

Acknowledgements

This research was funded by an ESRC Transforming Social Sciences collaboration grant (ES/N006461/1) to VL and SW, and by ESRC FRL Fellowship (ES/N017560/1) to SW. We thank Emma Bruce-Gardyne for proofreading the manuscript.

References

Agyei, S. B., van der Weel, F. R. R., & van der Meer, A. L. H. (2016). Longitudinal study of preterm and full-term infants: High-density EEG analyses of cortical activity in response to visual motion. *Neuropsychologia*, *84*, 89–104.

Anderson, A. J., & Perone, S. (2018). Developmental change in the resting state electroencephalogram: insights into cognition and the brain. *Brain and Cognition*, *126*, 40-52.

Azizollahi, H., Darbas, M., Diallo, M. M., El Badia, A., & Lohrengel, S. (2018). EEG in neonates: Forward modeling and sensitivity analysis with respect to variations of the conductivity. *Mathematical Biosciences & Engineering*, *15*, 905-932.

Babiloni, F., & Astolfi, L. (2014). Social neuroscience and hyperscanning techniques: past, present and future. *Neuroscience & Biobehavioral Reviews*, *44*, 76-93.

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

Baum, J. D., & Searls, D. (1971). Head shape and size of newborn infants. *Developmental Medicine & Child Neurology*, 13, 572-575.

Bell, M. A., & Cuevas, K. (2012). Using EEG to study cognitive development: issues and practices. *Journal of Cognition and Development*, 13: 3. Doi: 10.1080/15248372.2012.691143

Bell, M.A., & Wolfe, C.D. (2008). The use of the electroencephalogram in research on cognitive development. L.A. Schmidt, S.J. Segalowitz (Eds.), *Developmental Psychophysiology: Theory, Systems, and Methods*, pp. 150-170. Cambridge University Press, New York.

Béнар C.-G., & Gotman, J. (2002). Modeling of post-surgical brain and skull defects in the EEG inverse problem with the boundary element method. *Clinical Neurophysiology*, 113, 48–56.

Brooker, B. H., & Donald, M. W. (1980). Contribution of the speech musculature to apparent human EEG asymmetries prior to vocalization. *Brain and Language*, 9, 226–245.

Boto, E., Holmes, N., Leggett, J., Roberts, G., Shah, V., Meyer, S. S., ... Brookes, M. J. (2018). Moving magnetoencephalography towards real-world applications with a wearable system. *Nature*, 555, 657–661.

Burgess, A. P. (2013). On the interpretation of synchronisation in EEG hyperscanning studies: a cautionary note. *Frontiers in Human Neuroscience*, 7, 881. Doi: 10.3389/fnhum.2013.00881

Buzsaki, G. (2006). *Rhythms of the Brain*. Oxford University Press.

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

Clancy, R. R., & Legido, A. (1991). Postnatal epilepsy after EEG-confirmed neonatal seizures. *Epilepsia*, 32, 69-76.

Cohen, M. X. (2014). *Analyzing Neural Time Series Data: Theory and Practice*. MIT press.

Corrigall, K. A., & Trainor, L. J. (2014). Enculturation to musical pitch structure in young children : evidence from behavioral and electrophysiological methods. *Developmental Science*, 17, 142–158.

DeBoer, T., Scott, L. S., & Nelson, C. A. (2007). Methods for acquiring and analyzing infant event-related potentials. In M de Haan (Ed.) *Infant EEG and Event-Related Potentials*, pp. 5–37. New York, NY: Psychology Press.

De Haan, M. (2013). *Infant EEG and Event-Related Potentials*. New York, NY: Psychology Press.

Duc, G., & Largo, R. H. (1986). Anterior fontanel: size and closure in term and preterm infants. *Pediatrics*, 78, 904-908.

Fairhurst, M., & Dumas, G. (2019). Reciprocity and alignment: quantifying coupling in dynamic interactions. PsyArXiv, Preprint Doi: 10.31234/osf.io/nmg4x

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, 9602-9605.

Fisch, B. J., & Spehlmann, R. (1999). *Fisch and Spehlmann's EEG Primer: Basic Principles of Digital and Analog EEG*. Elsevier Health Sciences.

Folland, N. A., Butler, B. E., Payne, J. E., & Trainor, L. J. (2014). Cortical representations sensitive to the number of perceived auditory objects emerge between 2 and 4 months of age: electrophysiological evidence. *Journal of Cognitive Neuroscience*, 27, 1060-1067.

Friedrich, M., Weber, C., & Friederici, A. D. (2004). Electrophysiological evidence for delayed mismatch response in infants at-risk for specific language impairment. *Psychophysiology*, 41, 772-782.

Fujioka, T., Mourad, N., He, C., & Trainor, L. J. (2011). Comparison of artefact correction methods for infant EEG applied to extraction of event-related potential signals. *Clinical Neurophysiology*, 122, 43-51.

Georgieva, S., Lester, S., Yilmaz, M., Wass, S., & Leong, V. (2018). Topographical and spectral signatures of infant and adult movement artefacts in naturalistic EEG. *bioRxiv*, 206029. Doi: 10.1101/206029

Gibson, A., Bayford, R. H., & Holder, D. S. (2000). Two-dimensional finite element modelling of the neonatal head. *Physiological Measurement*, 21, 45-52.

Giraud, A. L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience*, 15, 511-517.

Gonçalves, S. I., de Munck, J. C., Verbunt, J. P., Bijma, F., Heethaar, R. M., & da Silva, F. L. (2003). In vivo measurement of the brain and skull resistivities using an EIT-based method and realistic models for the head. *IEEE Transactions on Biomedical Engineering*, 50, 754-767.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, 37, 424-438.

Grigg-Damberger, M., Gozal, D., Marcus, C. L., Quan, S. F., Rosen, C. L., Chervin, R. D., ... & Iber, C. (2007). The visual scoring of sleep and arousal in infants and children. *Journal of Clinical Sleep Medicine*, 3, 201-240.

Gwin, J. T., Gramann, K., Makeig, S., & Ferris, D. P. (2010). Removal of movement artefact from high-density EEG recorded during walking and running. *Journal of Neurophysiology*, 103, 3526-3534.

Haartsen, R., Jones, E. J., & Johnson, M. H. (2016). Human brain development over the early years. *Current Opinion in Behavioral Sciences*, *10*, 149-154.

Hagne, I. (1968). Development of the waking EEG in normal infants during the first year of life. In P. Kellaway and I. Petersen (Eds.) *Clinical Electroencephalography of Children*, pp. 97-118. Almqvist & Wiksell, Stockholm.

Hagne, I. (1972). Development of the sleep EEG in normal infants during the first year of life. *Acta Paediatrica*, *61*, 25-53.

Halit, H., Csibra, G., Volein, A., & Johnson, M. H. (2004). Face-sensitive cortical processing in early infancy. *Journal of Child Psychology and Psychiatry*, *45*, 1228-1234.

Hansman, C. F. (1966). Growth of interorbital distance and skull thickness as observed in roentgenographic measurements. *Radiology*, *86*, 87-96.

He, C., Hotson, L., & Trainor, L. J. (2007). Mismatch responses to pitch changes in early infancy. *Journal of Cognitive Neuroscience*, *19*, 878-892.

Hirasawa, K., Kurihara, M., & Konishi, Y. (2002). The relationship between mismatch negativity and arousal level. Can mismatch negativity be an index for evaluating the arousal level in infants? *Sleep Medicine*, *3*, S45-S48.

Hoehl, S., Michel, C., Reid, V. M., Parise, E., & Striano, T. (2014). Eye contact during live social interaction modulates infants' oscillatory brain activity. *Social Neuroscience*, *9*, 300-308.

Hoehl, S., & Wahl, S. (2012). Recording infant ERP data for cognitive research. *Developmental Neuropsychology*, *37*, 187-209.

Hyvärinen, A., & Oja, E. (1997). A fast fixed-point algorithm for independent component analysis. *Neural computation*, *9*, 1483-1492.

Jones, E. J., Venema, K., Lowy, R., Earl, R. K., & Webb, S. J. (2015). Developmental changes in infant brain activity during naturalistic social experiences. *Developmental Psychobiology*, *57*, 842-853.

King, J. R., Sitt, J. D., Faugeras, F., Rohaut, B., El Karoui, I., Cohen, L., ... & Dehaene, S. (2013). Information sharing in the brain indexes consciousness in noncommunicative patients. *Current Biology*, *23*, 1914-1919.

Kühn-Popp, N., Kristen, S., Paulus, M., Meinhardt, J., & Sodian, B. (2016). Left hemisphere EEG coherence in infancy predicts infant declarative pointing and preschool epistemic language. *Social Neuroscience*, *11*, 49-59.

Lachat, F., & George, N. (2012). Oscillatory brain correlates of live joint attention: a dual-EEG study. *Frontiers in Human Neuroscience*, *6*, 156. Doi: 10.3389/fnhum.2012.00156

Lachaux, J. P., Rodriguez, E., Martinerie, J., & Varela, F. J. (1999). Measuring phase synchrony in brain signals. *Human Brain Mapping*, 8, 194-208.

Legido, A., Clancy, R. R., Spitzer, A. R., & Finnegan, L. P. (1992). Electroencephalographic and behavioral-state studies in infants of cocaine-addicted mothers. *American Journal of Diseases of Children*, 146, 748-752.

Leong, V., Byrne, E., Clackson, K., Harte, N., Lam, S., & Wass, S. (2017). Speaker gaze changes information coupling between infant and adult brains. *Proceedings of the National Academy of Sciences*, 114, 13290-13295.

Leong, V., Noreika, V., Clackson, K., Georgieva, S., Brightman, L., Nutbrown, R., ... & Wass, S. (2019). Mother-infant interpersonal neural connectivity predicts infants' social learning. *PsyArXiv*, Preprint Doi: 10.31234/osf.io/gueaq

Leppänen, J. M., Peltola, M. J., Puura, K., Mäntymaa, M., Mononen, N., & Lehtimäki, T. (2011). Serotonin and early cognitive development: variation in the tryptophan hydroxylase 2 gene is associated with visual attention in 7-month-old infants. *Journal of Child Psychology and Psychiatry*, 52, 1144-1152.

Lew, S., Sliva, D. D., Choe, M. S., Grant, P. E., Okada, Y., Wolters, C. H., & Hämäläinen, M. S. (2013). Effects of sutures and fontanels on MEG and EEG source analysis in a realistic infant head model. *NeuroImage*, 76, 282-293.

Light, G. A., Williams, L. E., Minow, F., Sprock, J., Rissling, A., Sharp, R., ... & Braff, D. L. (2010). Electroencephalography (EEG) and event-related potentials (ERPs) with human participants. *Current Protocols in Neuroscience*, 52, 6-25.

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. Doi: 10.3389/fpsyg.2018.01862

Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique* (2nd edition). MIT press.

Lund, T. R., Sponheim, S. R., Iacono, W. G., & Clementz, B. A. (1995). Internal consistency reliability of resting EEG power spectra in schizophrenic and normal subjects. *Psychophysiology*, 32, 66-71.

Makeig, S., Bell, A. J., Jung, T. P., & Sejnowski, T. J. (1996). Independent component analysis of electroencephalographic data. *Advances in Neural Information Processing Systems*, 8, 145-151.

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, 1199-1208.

Martynova, O., Kirjavainen, J., & Cheour, M. (2003). Mismatch negativity and late discriminative negativity in sleeping human newborns. *Neuroscience Letters*, 340, 75-78.

Meltzer, L. J., Montgomery-Downs, H. E., Insana, S. P., & Walsh, C. M. (2012). Use of actigraphy for assessment in pediatric sleep research. *Sleep Medicine Reviews, 16*, 463-475.

Michel, C. M., & Brunet, D. (2019). EEG Source Imaging: a practical review of the analysis steps. *Frontiers in Neurology, 10*, 325. Doi: 10.3389/fneur.2019.00325

Miljković, N., Matić, V., Van Huffel, S., & Popović, M. B. (2010, September). Independent Component Analysis (ICA) methods for neonatal EEG artefact extraction: Sensitivity to variation of artefact properties. In *10th Symposium on Neural Network Applications in Electrical Engineering* (pp. 19-21). IEEE.

<https://ieeexplore.ieee.org/abstract/document/5644041>

Mourad, N., Reilly, J. P., de Bruin, H., Hasey, G., & MacCrimmon, D. (2007, April). A simple and fast algorithm for automatic suppression of high-amplitude artefacts in EEG data. In *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07* (pp. 393-396). IEEE. <https://ieeexplore.ieee.org/abstract/document/4217099>

Munsters, N. M., van Ravenswaaij, H., van den Boomen, C., & Kemner, C. (2019). Test-retest reliability of infant event related potentials evoked by faces. *Neuropsychologia, 126*, 20-26.

Nelson, C. A., & De Haan, M. (1996). Neural correlates of infants' visual responsiveness to facial expressions of emotion. *Developmental Psychobiology, 29*, 577-595.

Nunez, P. L., & Srinivasan, R. (2006). *Electric Fields of the Brain: The Neurophysics of EEG*. Oxford University Press.

Odabae, M., Tokariev, A., Layeghy, S., Mesbah, M., Colditz, P. B., Ramon, C., & Vanhatalo, S. (2014). Neonatal EEG at scalp is focal and implies high skull conductivity in realistic neonatal head models. *Neuroimage*, *96*, 73-80.

Ogilvie, R. D. (2001). The process of falling asleep. *Sleep Medicine Reviews*, *5*, 247-270.

Ollikainen, J. O., Vauhkonen, M., Karjalainen, P. A., & Kaipio, J. P. (1999). Effects of local skull inhomogeneities on EEG source estimation. *Medical Engineering & Physics*, *21*, 143-154.

Op de Beeck, H. & Nakatani, C. (2019). *Introduction to Human Neuroimaging*. Cambridge University Press.

Orehova, E. V., Stroganova, T. A., Posikera, I. N., & Elam, M. (2006). EEG theta rhythm in infants and preschool children. *Clinical Neurophysiology*, *117*, 1047-1062.

Otte, R. A., Winkler, I., Braeken, M. A. K. A., Stekelenburg, J. J., Van der Stelt, O., & Van den Bergh, B. R. H. (2013). Detecting violations of temporal regularities in waking and sleeping two-month-old infants. *Biological Psychology*, *92*, 315-322.

Parmelee Jr, A. H., Schulte, F. J., Akiyama, Y., Wenner, W. H., Schultz, M. A., & Stern, E. (1968). Maturation of EEG activity during sleep in premature infants.

Electroencephalography and Clinical Neurophysiology, 24, 319-329.

Pedroso, F. S., Rotta, N., Quintal, A., & Giordani, G. (2008). Evolution of anterior fontanel size in normal infants in the first year of life. *Journal of Child Neurology*, 23, 1419-1423.

Plöchl, M., Ossandón, J., & König, P. (2012). Combining EEG and eye tracking: identification, characterization, and correction of eye movement artefacts in electroencephalographic data. *Frontiers in Human Neuroscience*, 6: 278. Doi: 10.3389/fnhum.2012.00278

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, 180-187.

Popich, G. A., & Smith, D. W. (1972). Fontanels: range of normal size. *The Journal of Pediatrics*, 80, 749-752.

Reid, V. M., Striano, T., & Iacoboni, M. (2011). Neural correlates of dyadic interaction during infancy. *Developmental Cognitive Neuroscience*, 1, 124-130.

Reynolds, G. D., & Richards, J. E. (2009). Cortical source localisation of infant cognition. *Developmental Neuropsychology*, 34, 312-329.

Roche-Labarbe, N., Aarabi, A., Kongolo, G., Gondry-Jouet, C., Dümpelmann, M., Grebe, R., & Wallois, F. (2008). High-resolution electroencephalography and source localisation in neonates. *Human Brain Mapping, 29*, 167-176.

Salinsky, M. C., Oken, B. S., & Morehead, L. (1991). Test-retest reliability in EEG frequency analysis. *Electroencephalography and Clinical Neurophysiology, 79*, 382-392.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences, 36*, 393-414.

Shmueli, K., van Gelderen, P., de Zwart, J. A., Horovitz, S. G., Fukunaga, M., Jansma, J. M., & Duyn, J. H. (2007). Low-frequency fluctuations in the cardiac rate as a source of variance in the resting-state fMRI BOLD signal. *Neuroimage, 38*, 306-320.

Slugocki, C., & Trainor, L. J. (2014). Cortical indices of sound localisation mature monotonically in early infancy. *European Journal of Neuroscience, 40*, 3608-3619.

Smith, J. R. (1938a). The electroencephalogram during normal infancy and childhood: I. Rhythmic activities present in the neonate and their subsequent development. *The Pedagogical Seminary and Journal of Genetic Psychology, 53*, 431-453.

Smith, J. R. (1938b). The electroencephalogram during normal infancy and childhood: II. The nature of the growth of the alpha waves. *The Pedagogical Seminary and Journal of Genetic Psychology*, 53, 455-469.

Smith, J. R. (1938c). The electroencephalogram during normal infancy and childhood: III. Preliminary observations on the pattern sequence during sleep. *The Pedagogical Seminary and Journal of Genetic Psychology*, 53, 471-482.

Smith, J. R. (1939). The “occipital” and “pre-central” alpha rhythms during the first two years. *The Journal of Psychology*, 7, 223-226.

Souza, S. W., Ross, J., & Milner, R. D. (1976). Alterations in head shape of newborn infants after caesarean section or vaginal delivery. *Archives of Disease in Childhood*, 51, 624-627.

Striano, T., Reid, V. M., & Hoehl, S. (2006). Neural mechanisms of joint attention in infancy. *European Journal of Neuroscience*, 23, 2819-2823.

Stroganova, T. A., Orekhova, E. V., & Posikera, I. N. (1999). EEG alpha rhythm in infants. *Clinical Neurophysiology*, 110, 997-1012.

Trainor, L. J., Lee, K., & Bosnyak, D. J. (2011). Cortical plasticity in 4-month-old infants: specific effects of experience with musical timbres. *Brain Topography* 24, 192. Doi: 10.1007/s10548-011-0177-y

van der Velde, B., Haartsen, R., & Kemner, C. (2019). Test-retest reliability of EEG network characteristics in infants. *Brain and Behavior*, e01269. Doi: 10.1002/brb3.1269

Vigário, R. N. (1997). Extraction of ocular artefacts from EEG using independent component analysis. *Electroencephalography and Clinical Neurophysiology*, 103, 395-404.

Vinck, M., Oostenveld, R., Van Wingerden, M., Battaglia, F., & Pennartz, C. M. (2011). An improved index of phase-synchronisation for electrophysiological data in the presence of volume-conduction, noise and sample-size bias. *Neuroimage*, 55, 1548-1565.

Wass, S.V., Noreika, V., Georgieva, S., Clackson, K., Brightman, L., Nutbrown, R., Santamaria, L., & Leong, V. (2018). Parental neural responsivity to infants' visual attention: how mature brains influence immature brains during social interaction. *PloS Biology*, 16: e2006328. Doi: 10.1371/journal.pbio.2006328

Weickenmeier, J., Fischer, C., Carter, D., Kuhl, E., & Goriely, A. (2017). Dimensional, geometrical, and physical constraints in skull growth. *Physical Review Letters*, 118, 248101.

Weitzman, E. D., & Graziani, L. J. (1968). Maturation and topography of the auditory evoked response of the prematurely born infant. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, 1, 79-89.

Xie, W., & Richards, J. E. (2017). The relation between infant covert orienting, sustained attention and brain activity. *Brain Topography*, 30, 198-219.