Combining real-time analysis with machine learning to study neural correlates of social attention in infancy

https://eprints.bbk.ac.uk/id/eprint/53480/

Version: Full Version

Citation: Throm, Elena Veronika (2024) Combining real-time analysis with machine learning to study neural correlates of social attention in infancy. [Thesis] (Unpublished)

Combining Real-Time Analysis with Machine Learning to Study Neural Correlates of Social Attention in Infancy

Elena Veronika Throm

Supervised by

Professor Emily J. H. Jones

Professor Robert Leech

Thesis submitted for the degree of

Doctor of Philosophy, Ph.D.

Birkbeck, University of London

2024

Declaration

I, Elena V. Throm, hereby declare that the work presented in this thesis, with the exceptions laid out below, is entirely my own.

Chapters 3 and 4 in this thesis are based on analyses of longitudinal data that were collected as part of the British Autism Study of Infant Siblings (*BASIS-Phase-3 or "Studying Autism and ADHD in the eaRly yearS", STAARS*). This study was designed and the data collected and pre-processed by the STAARS team: Mary Agyapong, Tessel Bazelmans, Jannath Begum Ali, Tony Charman, Francesca Conti, Leila Dafner, Kim Davies, Mutluhan Ersoy, Janice Fernandes Zoë Freeman, Teodora Gliga, Amy Goodwin, Rianne Haartsen, Hanna Halkola, Alexandra Hendy, Rebecca Holman, Meg Jackson, Mark Johnson, Emily Jones, Sarah Kalwarowsky, Anna Kolesnik, Sarah Lloyd-Fox, Luke Mason, Nisha Narvekar, Laura Pirazzoli, Chloë Taylor and Natalie Vaz. I contributed with overall recruitment for the Babylab participant database and with taking care of EEG net maintenance and repair.

Chapters 5 and 6 are based on studies using Neuroadaptive Bayesian Optimisation (NBO) that combines real-time analysis of neurophysiological data with machine learning. The machine-learning scripts were created by Pedro F. da Costa. The real-time analysis scripts for Event-Related Potentials (ERP) used in Chapter 5 were created by Luke Mason and Rianne Haartsen and adapted and optimised for infant ERP and continuous EEG data by Anna Gui and me. Together with Anna Gui, I adapted and optimised the real-time analysis MATLAB scripts, designed the two studies, and implemented and piloted the NBO method with infants. I was responsible for recruitment, data collection, data processing and analysis of the two NBO studies in this thesis including EEG data and questionnaire data.

The following parts of this thesis are published or in preparation for publication:

The work presented in Chapter 3 and Chapter 4 is in preparation for submission for publication.

The work presented in Chapter 5 has been submitted for publication:
Throm, E.*, Gui, A.*, Haartsen, R., da Costa, P. F., Mason, L., Leech, R., & Jones E. J. H. (in prep.) Combining Real-Time Analysis with Machine Learning to Study Neural Responses to Familiar Faces during Infancy: A Proof of Principle Study

The work presented in Chapter 6 has been published:
Throm, E., Gui, A., Haartsen, R., da Costa, P. F., Leech, R., & Jones, E. J. (2023). Real-time monitoring of infant theta power during naturalistic social experiences. *Developmental Cognitive Neuroscience*, *63*, 101300.

Elena V. Throm
20 January 2024

Acknowledgements

This PhD work would not have been possible without my supervisors, Professor Emily Jones and Professor Robert Leech. I am extremely grateful for your exceptional availability and guidance, for providing challenges, pointing me towards new opportunities and sharing your passion. I am also deeply grateful to my examiners, Professor Sabine Hunnius and Dr Tomoki Arichi, who helped me improve this thesis a lot. I could not have undertaken this work without the support of the European Commission through the Marie Sklodowska-Curie Doctoral fellowship within the frame of the Horizon 2020 Program (grant agreement No. 642996). I would like to express my deepest appreciation to all the families who contributed their time to taking part in the studies presented in this thesis.

Many thanks to our wonderful BONDS team and my colleagues at the Babylab. Anna, Rianne and Pedro - thank you for your various guidance and support, for the fun and truly memorable hours of testing together, and for starting and never losing faith in our success in what has been an ambitious project. Teresa, Paola and Louisa – thanks so much for all your advice in various questions related to data acquisition and analysis. Many thanks to Abbey, for your help during your internship in the UK, to my PhD fellow Eirini and my SAPIENS fellows Dianna and Tanaya, to Marian and Francesca and to our placement and visiting students who all assisted in the studies: Charlotte, Elle, Ella, Amy, Rhiannon, Kitty, Maddie and Alice. Many thanks to Tamsin, Berta, Leslie, Sarela, Isabel and Annabel, for all your help with all things administrative.

I am deeply grateful to our wonderful SAPIENS consortium. To my fellow ESRs, with whom I went on this path together, exchanging and supporting each other. I enjoyed our training schools a lot (especially the in-person ones!). To our PIs who believed in our success and shared their personal experiences and perspectives with us, an invaluable resource. Special thanks to Robert Oostenveld for teaching me about streaming and real-time EEG data analysis.

I'd like to thank my parents, siblings, grandma, parents- and grandma-in-law, for all your support, not only but including your help with childcare during the pandemic. To my friends, many thanks for bearing with me over the years abroad, for your open ears and for celebrating with me the little successes on the way. Finally, I want to thank my own little family: for having been on this journey with me, your unconditional support and endless endurance.

Elena V. Throm
20 January 2024

Abstract

The period of infancy is characterised by rapid development in cognition and behaviour. Social brain development before social behaviour difficulties emerge is crucial for understanding trajectories and enabling early investigation. Investigating brain activity patterns in naturalistic contexts is particularly relevant for ecological validity and the generalisability of findings. Further, it is important to expand experimental tools to not only answer whether and how far brain responses differ in neurodivergent populations, but also what stimulation would be needed to optimally trigger social brain network activity in subgroups and individuals. The first aim of this thesis was to investigate whether brain activity during naturalistic social attention is altered in infants with an elevated likelihood for autism. The second aim was to apply a novel individualised experimental approach to identify what are the aspects of social interaction that optimally elicit social brain network activation in the individual infant. Chapter 1 covers theory and signs of social attention in infancy, with a focus on naturalistic contexts. Chapter 2 introduces the method Neuroadaptive Bayesian Optimisation and describes its application to infant encephalography data to reliably map an individual infant's neural response across a wide range of social stimuli. Chapter 3 shows that neural responses during social compared to non-social attention in a naturalistic setting are altered in infants with familial likelihood of autism, and Chapter 4 suggests that this difference emerges in the second half of the first year of age. Chapter 5 presents a proof-of-principle study testing the feasibility of Neuroadaptive Bayesian Optimisation with infant EEG data to study which among a range of familiar and unfamiliar faces maximally triggers social brain network activation in the individual infant. Chapter 6 expands this method to a naturalistic setting to identify which aspect of social interaction the individual infant maximally engages with. Implications of the findings and future research directions are discussed.

Table of Contents

List of Tables

List of Figures

CHAPTER 1: GENERAL INTRODUCTION

1.1 Introduction

As humans we are social beings. We are able to communicate our desires and beliefs, interpret mental states of other people, experience empathy, create long-lasting bonds, and coordinate complex global collaborations to achieve joint aims. When we are born, we depend for a relatively long time on the close and consistent care of people around us to ensure survival. Some theorists have argued that it is this long early period of close proximity with our caregivers in infancy due to our premature condition at birth that has been a critical factor in the evolution of humans as a high-functioning social species (Bjorklund & Green, 1992). Indeed, infants are surrounded by their caregivers and other people most of the time (Jayaraman, Fausey, & Smith, 2015). Social cognition develops rapidly during infancy, interactively shaped by genetic and environmental factors, and by the second year of age, children show impressive social interaction skills. Studies of decades of research have investigated how infants come to develop these skills, and which aspects in early development are linked to individual differences in childhood. This thesis contributes to this history of work by focussing on how infants process social information in a naturalistic context and testing the feasibility of a novel approach for studying infant social attention on the level of individuals and subgroups.

1.2 The concept of attention

Navigating in the complex social world requires infants to select from the myriad of options what aspects to pay attention to. Attention serves the selection of relevant information from the environment for further processing such as perception, action and memory, while inhibiting distracting, alternative information (Amso & Scerif, 2015; Scerif, 2010). According to this domain-general, functional account, attention is a set of processes that does not work in isolation but serves and depends on other processes such as motor control, arousal, higher-level goals, and prior knowledge.

An influential framework for attention in adults has been proposed by Posner and Petersen (Petersen & Posner, 2012; Posner & Petersen, 1990). This framework includes three components of attention which are supported by distinct neural networks and serve different functions related to the selection of relevant information and inhibition of distractors: 1) alerting for readiness or warning responses; 2) orienting for attention shifting; and 3) executive attention / executive control for switching, inhibiting, top-down control of visual attention.

Many theories have since built on the framework proposed by Posner and Petersen (1990), including theories of infant attention. For example, Colombo (2001) presented a theory for infant attention including very similar components as Posner and Peterson: alertness, spatial orienting and attention to object features, and endogenous/sustained attention (Colombo, 2001). A period of infant attention was thereby shown to consist of three phases, distinguished based on changes in heart-rate-measured arousal: 1) stimulus orienting, 2) sustained attention, and 3) attention termination (Richards, Reynolds, & Courage, 2010).

The development of the attention system is suggested to depend on connectivity between the hierarchically organised dorsal and ventral pathways and frontal cortical areas (Amso & Scerif, 2015). Thus, while visual attention in young infants initially may be involuntary and limited both on the level of sensory and motor functions as well as brain connectivity, over the first postnatal year, these connections are strengthened and especially in the second half of the first year increasingly linked to frontal brain areas, resulting in improving executive control functions and increasingly efficient selecting information and inhibiting distraction (Amso & Scerif, 2015; Johnson & De Haan, 2015).

1.3 Measuring attention in preverbal infants

In older children and adults, attention can be measured using tasks with explicit instructions for the participant. However, understanding and following these instructions is impossible for preverbal infants. Therefore, other measures are required to study attention in this population.

One way to investigate infant attention that has been used for decades to study infant attention and has revealed key insights into early social attention processes is by examining infants' looking behaviour. Looking behaviour can be video-coded manually after the task or automatically using eye-tracking technology based on infrared-light, for example.

In *habituation tasks*, looking time to a novel stimulus is measured after looking time to a previously presented stimulus has been established. Dishabituation, measured as an above-threshold-increase or recovery in looking time to the novel stimulus relative to the habituated stimulus, is being interpreted as successful discrimination of the novel from the habituated stimulus.

The *gap-overlap task* presents a central stimulus and, either simultaneously or after a brief gap, a peripheral stimulus, randomly on the left or right side of the screen. Attention disengagement is indexed by the reaction time in gaze shifting from the central to the peripheral stimulus. Using the gap-overlap task, for example, revealed prolonged reaction time of disengaging from the central stimulus to look at the peripheral stimulus in 14-month-old infants with later autism diagnosis compared to infants without later diagnosis (Elsabbagh, Fernandes, et al., 2013).

Another type of tasks that has been frequently used especially in the literature on infant social attention are *preferential looking tasks*. These tasks present to the infant a pair of images or a more complex scene and measure the time that the infant chooses to look at each image or at predefined areas of interest in the scene, for example, indexing attention orientation or executive attention. Preferential looking tasks studying attention modulation by social cues have revealed that newborns look longer at faces with direct versus averted gaze (Farroni, Menon, & Johnson, 2006; Farroni et al., 2002; Figure 1.1), and that looking at the eyes in dynamic videos was reduced while looking at the mouth was increased in 2-year old children with autism, compared to non-autistic children (Jones, Carr, & Klin, 2008). A more recent eye-tracking study observed preferential looking to the eyes versus mouth in videos of faces in 5-month-olds, across still and dynamic stimuli, and could show that this preference was influenced by genetic factors (Viktorsson et al., 2022).

Figure 1.1. Stimuli used in preferential looking task in Farroni et al., 2006, p. 300

Another task that has been used to study infant social attention is the *face-pop-out task*. It presents images of multiple objects, including a face, and measures initial orienting towards or engagement with the face compared to the distractor objects. Variables of interest in studying social attention can be whether or not the first look is towards the face, the latency of orienting towards the face stimulus, the length of the first face fixation, the return to the face after the initial fixation or the length of overall fixations. Studies using the face-pop-out task have, for instance, shown typical orienting responses towards a face (Elsabbagh et al., 2013) and shorter overall, but not differential, fixation durations (Wass et al., 2015) in infants with elevated likelihood for autism.

Imitation tasks have been used to measure whether infants are able to understand others' intentions. For example, infants have shown to more often imitate an adult's action if the action was labelled as intended visually at 12 months (Schwier, van Maanen, Carpenter & Tomasello, 2006) or verbally at 14-18 months (Carpenter, Akhtar & Tomasello, 1998).

Joint attention tasks measure whether the infant can jointly attend to an object or event with a social partner and at the same time understand that they do this from different perspectives and respond to or initiating events of joint attention through gaze and gestures (Tomasello, 1995). For example, establishing joint attention with an adult regarding a novel object significantly increases the subsequent looking time to the novel toy in 9-month-old infants (Striano, Chen, Cleveland & Bradshaw, 2006). Joint attention skills build the basis for developing theory of mind, the ability to relate others' behaviour to others' or one's own mental states such as intentions or emotions (Mundy, 2018).

While they have provided crucial insights into early development, looking measures do not tell us the full story. For example, infants were more attentive during heart-rate defined sustained attention versus attention termination, even though they were still looking at the stimulus (Richards, Reynolds & Courage, 2010), indicating that attentional states vary also without accompanying changes in looking behaviour. Besides gaze behaviour, recent advancement of more fine-grained eye-tracking technology has allowed to investigate also other measures of oculomotor activity, such as dilation of the pupil (Laeng, Sirois & Gredebäck, 2012). For example, pupil dilation upon listening to non-social versus sounds was greater in infants with versus without later autism diagnosis, and the extent of pupil dilation to non-social sounds was related to stronger autistic symptoms (Rudling, Nyström, Bölte, & Falck-Ytter, 2022). Other behavioural measures besides oculomotor activity include infant sucking rate that has been used to measure habituation to a stimulus, or infant heart rate to measure attentive states (Richards, Reynolds & Courage, 2010).

While behavioural measures have provided important insights into infant attention processes, they are limited to studying overt behaviour, while processing occurs not only on the level of overt behaviour (Richards, Reynolds & Courage, 2010). Further, behavioural tasks might make specific requirements to the infant (e.g. requiring certain motor activity) that are unrelated to and hence might cover the construct under study. Another point to consider is that processing differences might arise before they manifest in behaviour. Indeed, while infants with atypical social development showed typical proportions of looking time to the eyes in faces, processing alterations could be observed on the neural level (Elsabbagh et al., 2012). Therefore, neural activity can be a useful measure of infant attention.

Numerous studies investigating task-related neural activity in infants have used the method of *electroencephalography* (EEG; Richards, 2001). EEG directly measures the electrical activity of neurons in the cortex. It allows to detect rapid and subtle changes in neural processing and can thus help to study fast and brief cognitive processes. It can be used with a cap or net placed on the infant's head, allowing to some degree freedom of movement. As a disadvantage, it is limited to recording electrical activity in the cortex and does not provide the spatial resolution needed to investigate activation in deeper levels of the brain.

One aspect of the EEG signal that particularly makes advantage of the high temporal sensitivity inherent to EEG and that has been frequently used for investigating visual attention are *event-related potentials (ERPs)*. ERPs are changes in the EEG signal time-locked to a preceding cognitive event, such as the processing of a briefly presented stimulus (Luck, 2014). An ERP is obtained by averaging the signal across multiple equal trials in order to increase the signal-to-noise ratio and reveal the pure signal change specifically related to the event while cancelling out noise. One component of the infant ERP that has been related to attention engagement is the Negative central (Nc) component, a negative deflection of the ERP recorded over frontocentral regions of the cortex, peaking between 400 and 800 ms after stimulus onset (Haan, Johnson & Halit, 2003).

Besides ERPs, EEG can be used to study *neural oscillations,* that is synchronised rhythmic neural activity arising from regularly spiking populations of neurons. The EEG signal can be decomposed into different frequency bands reflecting the frequency of the rhythmic spiking. Stronger oscillatory power of one frequency band indicates stronger synchronised activity of source neurons in that respective frequency. The different frequency bands have been related to different functions (Saby & Marshall, 2012). For example, the theta band (with a frequency of 3-6Hz) has been related to information encoding and learning (Begus & Bonawitz, 2020; Jones et al., 2020; Klimesch, 1999), while the alpha band (with a frequency of 6-9Hz) has been related to inhibition control and memory performance (Klimesch, 1999; Orekhova, Stroganova, & Posikera, 2001). A more detailed description of the EEG method can be found in Chapter 2.

Apart from EEG, other infant-friendly, non-invasive techniques are neuroimaging techniques which allow for higher spatial resolution but at the downside are more limited in temporal resolution. *Functional magnetic resonance imaging (fMRI)* indirectly measures activity in subcortical structures by detecting changes in the blood oxygenation level affecting an artificially created magnetic field. Because fMRI requires participants to lie still in large scanners in order to not create motion artifacts, fMRI in developmental research has been restricted to studying sleeping infants (but see Ellis et al., 2020, for recent advances on fMRI with awake infants). Therefore, fMRI has mostly been used to study sleeping infants' brain responses to auditory stimuli.

Functional near-infrared spectroscopy (fNRIS) is a neuroimaging technique that can be used to study awake infants, because it is portable and light and can, as EEG, be attached to an infant-sized cap. It was developed more recently and has since been used in numerous studies to study social brain specialization in infancy. It uses near-infrared light to measure the activation of cortical regions by exploiting the facts that, first, active neurons consume oxygen and, second, that oxygenated haemoglobin (HbO2) in comparison to deoxygenated haemoglobin (HbR) in the blood absorbs more light and therefore reflects less. fNIRS shines infrared light from various source optodes through the scalp and measures by various detector optodes the amount that is reflected in a certain scalp location. While still being limited to measuring cortical activity, compared to EEG it provides a better spatial resolution of the activation across the cortex (Quaresima & Ferrari, 2019). Further, if different source-detector separations are used for the different source-detector pairs, fNIRS also allows information about activation in deeper cortical layers, with a depth sensitivity of about 1.5 cm (Quaresima & Ferrari, 2019). Unlike EEG, fNIRS is not measuring the activation of a group of neurons directly by recording the electrical potential that arrives at the scalp, but indirectly by inferring it from the hemodynamic response. Since this response takes multiple seconds to build up after the activation of a group of neurons, the temporal resolution of fNIRS is low compared to EEG and is best suited for block-design paradigms with stimulus presentation blocks of several seconds, allowing the response to build up and, interleaved by a baseline task of several seconds, return to baseline again.

Taken together, techniques measuring infant neural activity can capture variations in processing that are not visible on a behavioural level. Besides fMRI and fNIRS, EEG is a technique that has been widely used to study neural correlates of infant social attention. It is particularly useful for studying neural activity in awake infants because it is non-invasive, portable and light weight, thus tolerating a relatively high degree of movement while detecting rapid changes in the signal.

1.4 The concept of social attention – a developmental perspective

The term "social attention" has been used widely in the literature to refer to attention in the context of conspecifics. However, there is no agreement on how this term should be defined or measured (see Salley and Colombo, 2016, for a review). To characterize the existing developmental literature using the term social attention either conceptually or empirically, Salley and Colombo (2016) suggested a framework based on three separate but inter-related functions of social attention:

1. social attention behaviours (mainly joint attention) that can produce social interaction, typically developing by the age of 9 months (Mundy & Newell, 2009);
2. social motivation, referring to the degree to which attending to and/or engaging with the social world is experienced rewarding (Chevallier, Kohls, Troiani, Brodkin, & Schultz, 2012); and
3. basic visual attention to social information, for instance reflected in the tendency to orient towards social cues such as face-like stimuli and direct gaze from birth (Farroni, Csibra, Simion, & Johnson, 2002) which has been proposed to be fundamental for ensuring social experiences down the line (Jones & Klin, 2013; Klin, Shultz, & Jones, 2015).

It remains unclear whether social attention is a construct independent from other forms of (non-social) attention, or whether by contrast social developmental outcomes can be explained by general cognitive processes. It is further unknown whether social attention is a unitary construct, that is whether the three subconstructs of attention are or are not based on the same underlying processes.

Regarding whether social attention is a construct **independent** of non-social attention, Salley and Colombo (2016) demonstrate the possibility that in young infants, attention in social and non-social contexts is the same. Over the first year then, with the maturation of brain networks and the concurrent development of endogenous attention and executive function as well as the progressing cortical specialisation for social stimuli, social attention gradually becomes independent and is finally expressed in social interaction behaviours like joint attention. In other words, under this view, social and non-social attention are initially not distinguishable but diverge with increasing expertise and specialisation over the course of development. The dynamic cues in a social interactive context may thereby be what supports the development of social attention functions in unique ways. This proposal is supported by research showing stronger neural activation differences during social versus non-social processing in 12 versus 6 months olds (Jones et al., 2015), and an early attention bias towards social information observed already in newborns (e.g., Farroni et al., 2002) that may contribute to the development from basic attention to independent social attention behaviour. They call for more research explicitly comparing the two forms of attention longitudinally, since investigating attention including both social and non-social stimuli over multiple time points in early development is necessary to see whether results can be generalised to the other, respectively, and when. Including samples with both typical and atypical social development can thereby be particular helpful.

Regarding whether social attention is a **unitary** construct, Salley and Colombo (2016) suggest the three social attention functions may be distinct in younger infants but merge into a unitary social attention process over development that is finally manifested in joint attention behaviours. The final answer to the question of whether and when indeed the three functional components of social attention are first dissociated and over the course of development become the unified construct social attention requires systematically comparing each of them with multiple measures, respectively, in both social and non-social contexts across development (Salley & Colombo, 2016).

1.5 The development of the social attention network in the brain

Social attention in adults is proposed to be underpinned by activity in the social brain network, a set of different but functionally connected brain regions that that have shown to be involved in social attention processes. The social brain network includes the superior temporal sulcus (STS), the fusiform “face are” (FAA), and the orbito-frontal cortex (Adolphs, 2003; Klein, Shepherd & Platt, 2009). Various theories have been put forward aiming to explain the emergence of the social brain network. The core question has been to what extend and how over development brain and experience interact in fostering social learning in the individual.

Early theories taking a *maturational perspective* of development proposed that brain regions must become mature to allow for the emergence of a new behaviour or skill. This view predicted that the areas activated by social stimuli in children should be a subset of the areas activated by social stimuli in adults. Interregional connectivity does not play an intrinsic role in this account, while an increase in functional connectivity would be explained through the maturation of the relevant fibers (Johnson, Grossmann, & Kadosh, 2009).

On the other hand, theories taking a *skill-learning perspective* propose that changes in the brain happen based on experience and the acquisition of a new skill (e.g., Gauthier & Nelson, 2001). This view suggests that during social learning the same brain regions are active in infants as in adults and that social learning can occur at any point in life and is not bound to age or sensitive periods in development (Johnson, Grossmann, & Kadosh, 2009).

Finally, the *Interactive Specialization (IS)* framework has been proposed which combines both views (Johnson, 2001, 2011; Johnson et al., 2005). It focusses on the activation of cortical networks instead of isolated regions, that is the communication between anatomically different cortical regions. Response patterns in these networks change over development due to their interaction and competition with each other in the context of task processing. Thus, broad regions are initially activated by a wide range of tasks. With experience, anatomically distinct regions will become increasingly fine-tuned in responding to a subset of relevant tasks, such as seeing an upright compared to an averted face, and will become more interconnected in their specific responses, while at the same time responses within regions become more focal. Under the Interactive Specialisation framework, the social brain network emerges as a whole with development, becoming more fine-tuned to social versus non-social stimulation through competition of initially several involved regions activated by a broad range of social and non-social stimuli. Subcortical network activity from birth enables infants thereby to automatically and rapidly orient towards face-like stimuli, thus enabling increased exposure to social stimuli, initiating the tuning of brain networks towards social stimuli through experience (Johnson et al., 2005; Senju, Johnson, & Tomalski, 2014). This initial bias is taken over by the infants' ability to learn how to prioritise and use social information after the first few months after birth, resulting in progressive increase in sensitivity and selectivity to social versus non-social stimuli observed over the first year of life (e.g., Jones et al., 2015). Under this framework, early atypical processing of social stimuli can cause a subsequent lack of or altered specialization of the social brain network, with cascading consequences down the line, potentially giving rise to social behaviour difficulties in later in development. Therefore, studying brain responses to social versus non-social stimuli early on in individuals who go on to be diagnosed with a neurodevelopmental condition may reveal early deviations in social processing in these individuals and could facilitate an earlier diagnosis (Jones et al., 2017). Atypical social development is further discussed in Chapter 3.

1.6 Neural signs of infant social attention

1.6.1 Neural signs of infant attention to visual social cues

A substantial part of the literature into early social attention development has investigated infants' responses to visual social stimuli, primarily towards faces and face-like stimuli. Overall, findings support the view of an initial face bias and increasing cortical specialisation of the social brain network in the first year of age proposed by the Interactive Specialization framework.

Studies using looking measures showed that from 2 to 6 months of age, the time spent looking at the eyes increased in typical infants (Jones & Klin, 2013). At 6 months, the face was more often the first object infants looked at first among an array of objects in the face-pop-out task compared to the other objects; while this effect did not differentiate between upright and inverted faces it did differentiate between non-scrambled and scrambled faces (Gliga, Elsabbagh, Andravizou & Johnson, 2009). At 9 months, joint attention with an adult regarding a novel object significantly increased the subsequent looking time to the novel toy (Striano et al., 2006).

On a neural level, it was shown that shortly after birth, infants exhibited stronger neural correlates of attention to face-like stimuli compared to scrambled or inverted face-like patterns (e.g., Johnson, Dziurawiec, Ellis & Morton, 1991), as well as to open eyes and direct gaze compared to closed eyes and averted gaze (Farroni et al., 2002; Farroni, Johnson, & Csibra, 2004). Further, social brain activity in newborns was stronger in response to a dynamic face stimulus compared to a dynamic non-face stimulus (Farroni et al., 2013). Further, differential ERP responses have been observed to fearful and neutral faces in infants aged 3-6 months (Hoehl & Striano, 2010) and to faces displaying gaze shifts towards versus away from an object at 9 months (Atsushi Senju, Johnson & Csibra, 2006). Increasing specificity to upright versus inverted faces observed between 3 and 12 months (Halit, De Haan, & Johnson, 2003) suggests ongoing cortical specialisation to faces over the first year of age.

Infant ERP responses also reflected infants' recognition familiar faces on the level of the Negative central (Nc) component. The Nc has been linked to attention allocation (Conte, Richards, Guy, Xie & Roberts, 2020; De Haan, Johnson, & Halit, 2007; Richards, 2003). By 6 months, infants showed a differential Nc response mother's face versus a stranger's face (De Haan, 1997; De Haan & Nelson, 1999; Swingler, Sweet, & Carver, 2007; Webb, Long & Nelson, 2005).

Finally, cortical specialization may start from the moment the infant first experiences the world, as the magnitude of the strength in social brain activity during face versus non-face processing correlated with age in hours (Farroni et al., 2013).

1.6.2 Neural signs of infant attention to auditory social cues

Studies investigating infants' attention to auditory social cues have examined the mismatch negativity (MMN), a negative deflection of the ERP response to rarely occurring deviant tones, which has been used as a measure of memory of the standard tones (Cheour, Korpilahti, Martynova & Lang, 2001). The MMN has been shown to be modulated by speech-specific sounds in early infancy. For example, at 4 months, the topography of the mismatch negativity across the scalp differentiated between syllables and tones (Dehaene-Lambertz, 2000). The MMN rapidly develops in early infancy, as demonstrated by a greater response in newborns at older compared to younger gestational age (P. H. T. Leppänen et al., 2004), and by 6 months has reached its typical (negative) form for vowels but not consonants (Cheng et al., 2015), suggesting early specialisation towards speech over non-speech sounds.

Neural signs of social attention are not limited to speech but have also been investigated in the context of non-speech human sounds. As for speech-sounds, these studies have reported differential brain responses to social versus non-social processing. For example, by 7 months, infants showed differential responses to human sounds (e.g., footsteps, clapping, yawning, sighing) compared to nonhuman sounds (Blasi et al., 2011; Geangu, Quadrelli, Lewis, Macchi Cassia, & Turati, 2015), and selective responses for human versus nonhuman sounds became stronger between 4 and 7 months in typical infants (Lloyd-Fox, Blasi, Mercure, Elwell, & Johnson, 2012). Finally, responses were more focal in social brain network regions at 6 compared to 3 months (McDonald & Perdue, 2018).

These findings suggest an increase both in strength and spatial extent of differential responses to social versus non-social sounds across early infancy.

1.6.3 Neural signs of infant attention to audio-visual social cues

In daily life, infants experience social phenomena as a whole – embracing multiple modalities simultaneously as well as complex temporal dynamics. Findings from studies using more naturalistic, multimodal contexts support the proposed increase in specialisation. For example, infants showed differential theta power responses while watching a complex video with social compared to non-social information (Haartsen et al., 2022; Jones, Venema, Lowy, Earl & Webb, 2015), and the extent and strength of differential theta activity during an experimenter singing nursery rhymes versus operating a dynamic toy increased over the second half of the first year of age (e.g., Jones, Venema, Lowy, Earl, & Webb, 2015). Further there is evidence suggesting that entire networks seem to become more specialised towards social stimuli over this period. Differential connectivity of theta oscillations across brain regions during social versus non-social video viewing increased over the second half of the first year of age (van der Velde, White, & Kemner, 2021). Later in on in typical development, the preference for the face and the eyes seems to change with changing interest in the world. While viewing complex videos, the proportion of looking at the hands increased between 3 and 30 months, particularly when the hands performed an action (Frank, Vul, & Saxe, 2012). Further, an age-related shift was observed from looking to the eyes to looking to the mouth, especially when the face exposed emotional expressions or talking (Frank et al., 2012), suggesting that attention might be paid at the most relevant aspects at a certain point of development, respectively, which might be the eyes earlier in development and later the mouth, when language learning comes into play, as well as the face earlier in development and later the hands and actions, when infants start to engage in joint attention, predicting others intentions and following action sequences. Findings of these studies can provide important insights into what infants choose to attend to when presented with multimodal stimuli. In sum, multimodal dynamic stimuli allow studying infant attention in a more naturalistic context, by representing the complexity social cues are embedded in in the real world, and several studies have pointed towards progressing specialisation towards social versus non-social information presented in a naturalistic context.

Together, studies using various techniques and presenting stimuli of various modalities revealed signs of increased specialisation of brain activity towards social versus non-social stimulation over the first year of age, including in naturalistic settings with complex multimodal stimuli.

1.7 The relevance of naturalistic settings for studying infant social attention

While the majority of studies investigating infant social versus non-social attention used pre-recorded stimuli presented in isolation, to increase the generalisability of the findings to the real world, it has been acknowledged to also study responses to cues that combine various modalities and that are presented within complex dynamics, just as infants face them in daily life (Birmingham & Kingstone, 2009).

From an ecological validity perspective, stimuli should ideally reflect naturalistic social behaviour; at the same time, experimental control through standardisation is important in order to reduce the influence of confounding factors.
Live stimuli, as opposed to screen-based stimuli, involve a real person that functions as stimulus by displaying a certain behaviour or demonstrating an object, for example. Beyond the advantage of the increased ecological validity of findings from naturalistic settings, there is evidence suggesting that neural responses differ depending on whether social attention was studied using live or pre-recorded stimuli (Angelini et al., 2022; Jones et al., 2015). Specifically, live contexts showed to be more powerful

in eliciting selective responses to social stimulation: differential effects appeared more pronounced in a live social context involving a real social partner compared to in a screen-mediated context, suggesting enhanced processing of social cues when these are embedded in naturalistic settings (Jones et al., 2015; Striano, Reid, & Hoehl, 2006). Therefore, live contexts are particularly promising for the studying of also potentially small effects in the domain of social versus non-social attention.

But why do children perceive live stimuli differently? What the aspects are within social interaction that capture infants' interest? Theories bring forward different arguments of what makes social interaction special for us and *why* infants may prefer live stimuli over "artificial" ones. One critical element in live social interaction have been suggested to be ostensive cues, which are the aspects of human behaviour that signal to the perceiver that they are being addressed (G. Csibra & Gergely, 2009). The theory of *infant natural pedagogy* (Gergely Csibra & Gergely, 2006) proposes that infants attend to social stimuli because they are born with a "well-organised package of biases, tendencies and skills" (Csibra & Gergely, 2006, p. 8) making them prone towards adults who intentionally communicate with them in order to convey a piece of information. According to this account, eye contact, contingency and infant-directed speech are ostensive or communicative cues which infants have an innate tendency to react to. Ostensive cues signal to the infant that an interaction partner is addressing them and is about to transfer knowledge to them (Csibra & Gergely, 2009; Senju & Csibra, 2008). Thus, attending to ostensive cues helps infants learn the information an adult deliberately wants to convey to them. According to this account, for example, infants follow gaze because the other's gaze indicates that they are being addressed, and they only do so if accompanied by at least one of a set of specific ostensive cues (direct gaze, infant-directed speech, or contingent responsivity; Csibra, 2010). This view is in line with the findings of stronger social brain activation during overall social/communicative versus non-social/non-communicative situations (Blasi et al., 2011; Jones et al., 2016; Jones et al., 2015; Lloyd-Fox et al., 2013). Further specific support comes from eye tracking research. One study showed that both ostensive signals (IDS) and mere attention grabbers (e.g., beep sound) elicited gaze following in the infant, but referential object learning only occurred when preceded by ostensive signals, not by mere attention grabbers (Okumura, Kanakogi, Kobayashi,

& Itakura, 2020). This finding suggests that it is ostension, the deliberate intention to convey information and the hint towards it, beyond merely grabbing the child's attention, that is relevant in social learning (Okumura, Kanakogi, Kobayashi & Itakura, 2020). On the other hand, there is evidence challenging the view of the natural infant pedagogy theory. For example, in an eye tracking study, infants showed gaze following even without communicational content (Gredebäck et al., 2018).

Multimodal social stimuli might according to the infant natural pedagogy theory be particularly powerful as they combine multiple ostensive cues. While also videos of adults can be considered to convey a combination ostensive information, the mere finding oneself within the same context as the other person and being aware that the other person is a potential interaction partner might act as additional ostensive cues in a live compared to a screen-mediated context.

Minagawa, Xu & Morimoto (2018) have suggested other aspects that might explain why children perceive live paradigms differently, and these might function as additional ostensive cues enhancing in the infant their feeling of being addressed.
One property inherent to live but not screen-based stimuli are the enhanced sensory and perceptual characteristics, including increased size, three-dimensional information, and haptic and olfactory information (Minagawa, Xu, & Morimoto, 2018). A further factor to consider is the mere human presence (i.e. "a presence of a mind that encompasses intention and emotion", Minagawa, Xu & Morimoto, 2018, p. 4). Indeed, social psychology studies have observed the mere presence of a human to alter task performance ("social facilitation effect" and "social pressure"; Bond, C. F., & Titus, L. J., 1983).

Of note, while live studies can be either non-interactive or interactive, the present thesis focusses on non-interactive live studies. Live studies that additionally include an interactive component arguably further increase ecological validity but also further reduce experimental control. The additional interactive component additionally implies contingency (responsiveness) (Minagawa, Xu & Morimoto, 2018).

1.8 Theta power as a measure of infant attention in naturalistic contexts

1.8.1 Theta power as a measure of infant attention

One of the measures that has been used in the studies on infant attention in live social contexts are theta oscillations, reflecting slow frequency synchronised neural activity recorded at the scalp using electroencephalography (EEG). A detailed description of the method of EEG and theta oscillations can be found in Chapter 2.

The adult theta band (4-8 Hz) has shown to be related to cognitive processes involved in attention engagement and working memory (Cavanagh & Frank, 2014; Klimesch, 1999; Sauseng, Griesmayr, Freunberger & Klimesch, 2010). For example, in older children and adults, theta power was enhanced during social exclusion (van Noordt, White, Wu, Mayes, & Crowley, 2015) as well as during negative social feedback (van der Molen, Dekkers, Westenberg, van der Veen, & van der Molen, 2017). Similarly, older children and adolescents showed greater theta power to social rejection compared to acceptance by a peer, while adults showed greater theta power to threads of exclusion compared to social rejection and acceptance (Tang, Lahat, Crowley, Wu, & Schmidt, 2019).

In infants, the theta band oscillates with a 3-6 Hz frequency[1]. Infant theta has been associated processes related to learning. Infant theta power has been linked to attention engagement and information encoding (Anderson, Perone & Gartstein, 2022; Angelini et al., 2022; Bazhenova, Stroganova, Doussard-Roosevelt, Posikera & Porges, 2007; Begus, Gliga & Southgate, 2016; Jones, Venema, Earl, Lowy & Webb, 2017; Jones et al., 2015; Orekhova, Stroganova, Posikera & Elam, 2006; Smith et al., 2021; Stroganova, Orekhova & Posikera, 1998; for a review see Begus & Bonawitz, 2020). Specifically, infant theta may be involved in infants' updating of the emerging concepts they have of their environment after violation of their expectations (Berger & Posner, 2022). As such, in 9-month-olds, theta power increased shortly after the presentation of unexpected but not after expected pictures (Köster, Langeloh, Michel

---

[1] Recently, the definition has been adapted to 2-5 Hz (Xie, Mallin, & Richards, 2018). However, this thesis and the studies presented still use the 3-6 Hz definition in order to be consistent with the previous studies upon which this work builds.

& Hoehl, 2020), and under an incongruent versus congruent gaze shift live situation (Angelini et al., 2022). Further, theta power might be involved in attention. In 12-month-olds, theta activity during playing with an object predicted the length of infants' subsequent visual attention, while the prediction was weaker when the infants' visual attention was guided by an adult interaction partner (Wass et al., 2018). In 6-month-olds (Braithwaite, Jones, Johnson, & Holmboe, 2020) and 12-month-olds (Jones et al., 2020), there was an increase in theta power over the frontal cortex while infants were watching novel videos, and at 12 months, the amount of this increase was associated with neural responses while watching the same video again, reflecting the infants' learning of that video at 12 months. Crucially, these the differences in amount of learning the video correlated with cognitive skills at 9 months in typical infants (Braithwaite et al., 2020) as well as at 1, 2, 3 and 7 years in infants with older siblings with autism (Jones et al., 2020). These findings suggest that task-related theta power modulation reflects attention engagement and learning processes in infancy, and that the magnitude of this modulation is able to predict long-time cognitive performance.

Taken together, the current literature suggests that elevated theta power in infants may reflect an elevated attentional state in which infants are optimally set to receive information that helps them learn about the world.

1.8.2 Differential theta power as a measure of social versus non-social infant attention in naturalistic settings

A series of studies have investigated theta power responses to naturalistic social and non-social stimuli. Investigating theta power responses to videos either showing women singing nursery rhymes with gestures or showing dynamic toys (e.g., a rolling ball) revealed that theta power was stronger during viewing social versus non-social videos at 14 months (Haartsen et al., 2022) and at 12 months (Jones, Venema, Lowy, Earl & Webb, 2015; Jones et al., 2020), but not at 6 months of age (Jones, Venema, Lowy, Earl & Webb, 2015; Jones et al., 2020).

Further, the increase in differential theta power between 6 and 12 months was greater for social compared to non-social videos (Jones et al., 2017; Jones, Venema, Lowy, Earl & Webb, 2015).

In live stimulation contexts, elevated theta power towards social versus non-social stimulation was observed already earlier in infancy, at 6 months, and showed to increase in strength and spatial extent over the second half of the first year of age (Jones et al., 2015), suggesting that live social contexts are particularly powerful in triggering social brain responses at a young age. At 12 months, the effect appeared more pronounced in the live than in the screen context (Jones et al., 2015). Both at 6 and 12 months, the effect was strongest over frontal electrodes, but only at 12 months only it was also present over parietal and occipital regions, indicating an increase in topographical extent of differential theta activation during social versus non-social stimulation over the second half of the first year of age. While a differential response was also found in the alpha band, with greater alpha suppression during social versus non-social attention, this effect did not change with age (Jones et al., 2015), suggesting theta power differences to be a more sensitive measure of social brain specialisation in live social contexts. The effect of stronger theta power over frontal electrodes during live social versus non-social action in 6- and 12-month-old infants was replicated in an independent sample (Jones, Venema, Earl, Lowy & Webb, 2017).

Differential theta power responses to social live action might be related to individual differences in development. In fact, the stimulus effect showed to be reduced in infants of parents with higher social discomfort, avoidance and distress, both at 6 and 12 months (Jones, Venema, Earl, Lowy & Webb, 2017). Reduced parental social motivation was also related to other measures linked to social attention (shorter peak look duration to faces at 6 and 12 months, reduced P400 response to faces versus objects at 6 months; Jones, Venema, Earl, Lowy & Webb, 2017). This finding suggests a relation between reduced social brain specialisation and lower parental social motivation.

Infant theta power may further be increased during interaction-initiating behaviour in infants. For example, theta power (here: 3.6-5.6 Hz) was stronger during looking at a live neutral face compared to a smiling face in 5-month-old infants, accompanied by an increase in positive affect (Bazhenova et al., 2007). The authors suggested this pattern to reflect increased engagement with the neutral stimulus to achieve re-eliciting a response in the interaction partner. As in the study by Jones and colleagues

(2015), the effect was observed over anterior regions, but also in right posterior temporal regions.

Within different social live contexts, attention might be enhanced in the presence of a more engaging, addressing social partner offering more opportunities for learning. For example, frontal and temporal theta power was enhanced while an experimenter was labelling and demonstrating an object versus not labelling and merely handling the object in 11-month-old infants (Begus, Gliga & Southgate, 2016). Similarly, infants responded with stronger bilateral temporal (but, interestingly, not frontal) theta power in expectation of a person speaking in their mother-tongue versus foreign language, presumably expecting information from the former to be more relevant to them and bear greater potential for learning than the latter, at an age where word learning in the language of the mother-tongue is about to start. Together, the authors interpreted that the effect might be present over frontal regions in anticipation of receiving *functional* information, and over temporal regions in anticipation of *verbal* information (Begus, Gliga & Southgate, 2016). Recently, theta power has been studied during free play of infants with their caregivers. Stronger theta power was observed during infants' exploration of an object in interaction with their mother relative to watching dynamic videos in 6-to-12-month-olds (Anderson et al., 2022) which is in line with the previous findings of stronger theta power both for social versus non-social processing and live versus screen contexts. Further, during free play, theta power was observed to be stronger during interaction with the mother compared to looking at objects in the room in 6-month-olds, especially over temporal electrodes (Smith et al., 2021), indicating deeper attention engagement during interaction. Of note, in this study theta power was also stronger during independent play compared to looking at objects, contrary to the predictions of the authors. They argued that this result might be linked to characteristics of the analysis that was required given the (otherwise desirable) non-standardised nature of the free play paradigm. Specifically, behaviours were coded into event categories, potentially resulting in great behavioural variability within the categories in both mother-child interaction and independent play.

Together, there is strong evidence that theta power is elevated during social versus non-social stimulation, and this effect appears earlier in development and stronger in a live compared to screen context.

These findings also suggest that there are differences in how the effect of stronger social versus non-social theta power is distributed over scalp, while there is no clear consensus in the findings as to which regions exhibit elevated theta power to social stimulation at what age. According to the Interactive Specialisation hypothesis, responses become more specialised towards a restricted set of stimuli over development, and at the same time anatomically distinct regions specialising to this type of stimulation connecting with each other, reflecting a greater extent of specialised responses across the brain with development.

Besides changing with age, the location of the effect might depend on the type of information that is processed, for functional information frontally and for verbal information temporally (Begus, Gliga & Southgate, 2016). Further, a region effect has been suggested to vary depending the exact frequency range within the theta band, with an effect in frontal regions being specific to the lower theta range (4.0 and 4.4 Hz) and in temporo-parieto-occipital regions to the higher frequency range (5.6 Hz) (Orekhova, Stroganova, Posikera, & Elam, 2006). Yet other findings suggest no region differences at all (Stroganova, Orekhova & Posikera, 1998). Overall, it seems that the localisation of the effect depends on age, and possibly on the domain of processing and the exact frequency range under study.

In sum, research into infant theta power during naturalistic experiences indicates that theta power is enhanced during attention to a social partner versus an object, this differential response becomes stronger over the second half of the first year of age, reflecting increased specialization to social information, and naturalistic contexts may enhance the social/non-social effect on theta power. Theta power can hence be considered as a promising measure to study attention engagement during live social experiences in infancy.

1.8.3 What does theta power represent on a brain processing and neurobiological level?

Through coordinated neural activity, multiple bran regions can form brain networks to coordinate their activity (Klimesch, 1999). Theta rhythms, oscillating at a frequency of 3-6 Hz (or 2-5Hz; Xie, Mallin, & Richards, 2018) in the infant brain, result from joint activity of multiple neural populations that are firing or being inhibited together upon repetitive stimulation (Klimesch et al., 1999). The cognitive function that has been related to the theta rhythm in infants are information encoding and attentional control involved in active learning and memory formation. It is thought that the mechanism through which theta activity during information encoding is beneficial for learning is by affecting synaptic plasticity fostering inter-neural information transmission and hence across different regions of the brain (Buzsaki, 2006). Specifically, theta rhythms are thought to support information transmission at the synapse by enhancing dendritic long-term potentials (Rutishauser, Ross, Mamelak, & Schuman, 2010). Long-term potentiation (LTP) represents synaptic plasticity that facilitates signal transmission across neurons and has shown to facilitate memory and learning. Indeed, LTPs more likely arose when incoming stimulus sequences in the hippocampus happened at the theta frequency compared to other frequencies (Greenstein et al., 1988), and the joint firing of hippocampal neurons at the theta-rhythm during the active encoding of new information was predictive of memory performance (Rutishauser et al. (2010). Frontal theta power also influenced theta activity in the nucleus accumbens, and even more when a high reward was anticipated (Cohen et al., 2012) suggesting theta power being involved in reward-processing. It has been suggested that theta activity at the cortex has its source in cortical-hippocampal feedback loops (Klimesch, 1999). Indeed, phase synchrony between the cortex and hippocampus within the theta frequency band suggested theta oscillations to be involved in the communication between cortex and hippocampus (Lega, Jacobs, & Kahana, 2012).

In sum, enhanced theta power during social versus nonsocial attention might foster active learning and social reward processing or increased attentional control in the light of the combination of cues that need to be integrated in complex contexts, by enhancing information processing at the synapse.

1.9 Summary

The Interactive Specialisation framework of social brain development proposes an initial orientation bias towards face-like stimuli at birth, followed -with the onset of experiences in the world- by an increasing cortical specialisation towards social information over the course of infancy. With progressing cortical specialisation towards social stimuli over infancy, infant social attention may diverge as an independent concept separate from non-social attention. It serves three different functions related to the selection of information and the inhibition of distractors in the social world: social behaviour, social motivation, and basic visual attention, which may merge over infancy and consolidate as joint attention behaviour.

Infant social attention has been investigated using overt looking paradigms or neural measures. EEG is particularly suited for studying infant visual attention given its light weight and high temporal resolution and has been widely used to study infant social attention, primarily towards isolated visual stimuli presented on a screen. Findings underpin a preference for face-like versus non-face-like stimuli from birth, and an increase in the strength and extent of socially specific responses over the first year of age. The theta frequency band of the EEG signal has been related to attention engagement and information encoding and was greater during social versus non-social attention in a naturalistic context from 6 months of age, with an increasing differential response over the second half of the first year. Naturalistic contexts hence not only increase ecological validity of the findings and but also seem to enable enhanced processing of social cues and hence be particularly powerful at eliciting differential responses to social versus nonsocial information.

1.10 The present thesis

This thesis aims to both contribute to understanding early social development and to advance the tools for its investigation. It does so, first, by comparing theta activity during live social and non-social processing between infants with typical and atypical social development, to assess potential early processing alterations in a naturalistic context. Second, it introduces and discusses a novel experimental approach for identifying in the individual infant the optimal stimulus from a range of social cues that maximally triggers selected neural correlates of attention in the individual infant.

Chapter 2 describes the method of EEG and the application of a novel experimental approach to identify an individual infant's optimum social stimulus from a range of possibilities. Chapter 3 tests whether theta power during naturalistic social versus non-social information processing is altered 14-months-old infants with familial likelihood of autism. Chapter 4 investigates whether this alteration is already present in early infancy. Chapter 5 presents a proof-of-principle study testing the feasibility of the individualised approach with infant neurophysiological data. Chapter 6 applies the NBO approach to identifying individual infants' optimum stimulus from naturalistic social experiences. The findings of the present thesis are summarised, and implications and limitations of the present work as well as endeavours for future research are discussed.

CHAPTER 2: NEUROADAPTIVE BAYESIAN OPTIMISATION WITH INFANT ENCEPHALOGRAPHY DATA

2.1 Introduction

Traditional studies in developmental neuroscience typically explore how a stimulus evokes activity in a brain network. However, these findings do not answer what this activity means, since only the relation with one stimulus has been investigated. Likewise, it could also be evoked by other tasks. Further, there might be individual differences in the profile of which tasks trigger activity in a brain network. The standard experimental approach can only tackle these questions by the costly running of multiple consecutive experiments, with countless infants.

The present chapter introduces the novel experimental approach of "Neuroadaptive Bayesian Optimisation" to neurodevelopmental research, which maps within one brief experimental session how various stimuli differentially trigger social brain activity in the individual infant and can be used complementary to the traditional group-level approach to answer questions on level of individuals or subgroup. Its particular advantages for the field are described as well as the different components of the method in light of their application to studying brain function in infants.

Of note, while this methods chapter focusses using the novel experimental approach of Neuroadaptive Bayesian Optimisation with infant neurophysiological data, studies applying this new method are presented later in this thesis (Chapters 5 and 6), while Chapters 3 and 4 use the traditional experimental approach.

2.2 The limits of the classic experimental approach for studying variability in infant brain function in heterogeneous groups

Traditional studies require the investigation of average responses to few pre-selected stimuli. While this is necessary to draw robust group-level conclusions about a circumscribed effect, it limits the question we can ask about the process under study. Were the pre-selected stimuli the most relevant stimuli to study a phenomenon? What would responses look like across other stimuli that could not presented within the same session due to practical constrains? Addressing these questions would require multiple sequential experiments of the same kind, each investigating another small subset of stimuli of interest. This requires many resources and is in sum *inefficient* and decelerating the scientific progress. Are there subgroup in the sample that show different results? Addressing this question is not possible when responses are averaged across the sample in order to cancel out noise.

There are further problems related to the restriction to few stimuli and the pre-selection of these: Trying to preselect the one or two stimuli that reveal differences between individuals or different cultural subsets implies the *risk of missing crucial stimuli* by getting the preselection wrong. Further, preselection by a researcher prohibits studying which among those cues do individuals would come to choose to attend if they had the choice. Finally, being restricted to few stimuli often results in selecting the stimuli that are preferred by the *typical population* and in studying how atypical populations differ in the magnitude or location of their responses towards these normative stimuli. However, under a neurodiversity perspective (Pellicano & den Houting, 2022), it would be more inclusive to ask about *what stimuli* are preferred by different subgroups. Designs targeting this type of questions would not only ask more inclusive questions, but also produce results that could inform tailored interventions based on identified individual strengths rather than aiming to address common weaknesses.

Besides the restricted number of stimuli and their pre-selection, the embedded analytic flexibility in analysis after the data is collected bears the risk of hawking and p-hacking, ultimately resulting in a large part of the findings not being replicable (Ioannidis et al., 2014). Preregistration cannot fully solve the problem either (Nosek, Ebersole, DeHaven, & Mellor, 2018).

These limitations inherent to the traditional design are even more pronounced when studying infants. First, compared to research with adults, the field of developmental neuroscience is still relatively young. Hence, the preselection of stimuli is based on fewer accumulated knowledge than in the adult literature. This enhances the risk of getting the stimulus-selection wrong by missing important stimuli that have not been investigated yet. Second, infant research is progressing slower, given the more dynamic and flexibility-requiring nature of working with its subjects, as well as the smaller population size available for participation in studies. Therefore, it would benefit even more from overcoming the inefficiency embedded in traditional approaches in order to answer the questions at hand. Third, in infancy, the heterogeneity in neural processing between individuals is multiplied by the heterogeneity of the developmental change between individuals. Thus, a more efficient design would particular benefit longitudinal studies when investigating heterogeneous groups. Finally, flexibility in the choice of analytic parameters is even greater in a population for which there is little consensus on analytic standards.

In sum, the traditional experimental approach is restricted in the number of stimuli and relies on their pre-selection by the researcher and allows analytic flexibility that may contribute to low replicability of the findings. These restrictions can cause problems to the interpretation of the findings and at best limit the questions that can be asked.

2.3 A Novel Experimental Approach: Neuroadaptive Bayesian Optimisation

Recent advances have enabled the development of an individualised experimental approach that is able to complement the traditional group approach. Neuroadaptive Bayesian Optimisation (NBO, Lorenz et al., 2018) flips the rationale of the classic approach, in which an experiment aims to identify how individuals vary in magnitude or location of their responses to few pre-selected stimuli, towards mapping how a pre-defined brain function in the individual varies across a range of stimuli, revealing which are the stimuli that more or less trigger a brain response of interest in this individual. NBO uses a closed-loop design. It presents a stimulus selected from the range of prepared stimuli, analyses the individual's response to that stimulus, and based on this response selects the next stimulus, iteratively building up a model of the individual's response function across the stimulus space. Hence, an NBO experiment consists of two main processes: 1. Collecting and analysing neurophysiological data in real time and 2. Using the algorithm to build a model of the unknown brain function based on which the next stimulus is being selected (neuroadaptive). These two processes take turns in an iterative process working towards predicting the unknown underlying brain response function across one or more stimulus dimensions, until reaching a stopping criterion. The stimuli are created and arranged along the respective dimension(s) prior to the session.

The advantage of NBO is that it tests multiple predictions at the same time, by mapping responses across a wide stimulus space, without the need of presenting each single stimulus. Instead, useful stimuli are selected for presentation taking into account where in the stimulus space are predicted maxima and the uncertainty of the predicted values. In the closed-loop design, further points are sampled iteratively, until the algorithm has built a model of the unknown response function across the space based on the empirical samples it took of selected stimuli in the space.

NBO in principle works with any type of true, underlying relationship between the stimuli and the target metric (e.g., linear function of brain response changing from one extreme towards the other, U-shaped function), because it is non-parametric; it does not make prior assumptions about the underlying function and instead operates fully data-driven, that is only on basis of the empirical data it samples as measurements of the unknown, objective function.

The Neuroadaptive Bayesian Optimisation approach has been developed and validated in a proof-of-principle fMRI study with adults, by confirming with the individualised approach strong hypotheses about which visual and auditory properties best evoke a certain target brain state (Lorenz et al., 2016). This upside-down approach of experimental testing since allowed new endeavours into testing the meaning of brain function in adults, by mapping responses across a range of stimuli. At the same time, it allows broader hypotheses and a design that is pre-registered by nature, fostering reproducibility of research findings (Lorenz, Hampshire, & Leech, 2017), and due to its efficacy and robustness has particular value for the relatively young field of developmental neuroscience (Gui et al., 2022).

The approach has since been applied to studying the involvement of overlapping brain networks in various cognitive tasks (Lorenz et al., 2018). Further, it has been used to identify brain stimulation parameters that trigger phosphenes in healthy adults (Lorenz et al., 2019), and to identify in stroke patients the tasks by which their domain-general fronto-parietal networks is still activated, in order to inform effective rehabilitation protocols based on the individual neural profile (Lorenz et al., 2021). Recently, new methods for artifact rejection specifically tailored to the neuroadaptive approach have been developed (Ouyang, Dien, & Lorenz, 2022).

The individualised NBO approach can complement the traditional approach studying robust effects between stimuli on a group level, together maximising the number of research questions that can be asked. While the traditional approach asks, how on the group level brain responses are modulated by few preselected stimuli, NBO asks which of a range of possible stimuli is optimal for a subgroup or an individual to reach a specific target brain state. Not only can the two approaches study a subject from different perspectives, but they can also mutually inform each other: The group approach informs the NBO approach, for example, about what are robust and clinically relevant target brain metrics to optimise in an NBO experiment, and the NBO approach can inform the group approach, for example, about which stimuli are relevant for different subgroups.

2.4 The Method of Electroencephalography (EEG)

The cortical source of the EEG signal. The EEG signal measures voltage changes in the brain over time. During activation of a neuron, the balance of the voltage inside versus outside that neuron changes. Currents in the extracellular field respond to this voltage change, and together with currents in the extracellular field of neighbouring neurons create a local field potential (LFP). This LFP can be measured at the scalp by EEG sensors. Because the measured signal arises from different sources, and different sensors may measure activity from the same source, additional methods on the level of mathematics are needed to localise the source of the recorded EEG signal (Conte, Richards, Guy, Xie & Roberts, 2020).

EEG analysis. Before the EEG signal is being analysed, it is being cleaned or “pre-processed”. This step, for example, removes potential drifts and movement artifacts from the signal. The pre-processed signal enters analysis. Different aspects of the EEG signal have been analysed in relation to cognitive processing. Two major analysis methods of the EEG signal include *Event-Related Potentials (ERPs)* and *frequency analysis* of ongoing oscillations.

ERPs are direct rapid changes in the EEG signal time-locked to a preceding cognitive event, such as the processing of a briefly presented stimulus (Luck, 2014). This allows to relate measured voltage changes to a specific, e.g., cognitive, function. The overall EEG signal reflects a combination of activity. The ERP is obtained by averaging across multiple equal trials and reveal the pure signal change specifically related to the event while cancelling out noise. In principle, the more trials are averaged, the higher is the signal-to-noise ratio. Of note, the background noise affecting the signal-to-noise ratio varies between measurements and individuals. The different positive and negative deflections in the ERP waveform are usually named after the direction of the deflection (positive – P or negative – N) and their temporal position in the waveform (1, 2, 3 etc.) or their latency relative to stimulus onset (e.g. 100 ms). Other components have names relating to their deflection and position at the scalp, e.g., the Negative central component, Nc.

Components of the *ongoing oscillations* of the EEG signal can be studied, too. Changes due to an event can, for example, be reflected by a change in power of the oscillations. The ongoing EEG signal can be decomposed into frequency bands, that is waveforms oscillating in a certain frequency, measured in Hertz (Hz), reflecting how many oscillatory cycles per second the signal includes. The following frequency bands have been defined in relation to different functions in humans: Delta (0.5-2 Hz) – sleep, Theta (4-7 Hz) – attention engagement and learning (Orekhova et al., 2006), Alpha (9-11 Hz) – attentional control, inhibition, task engagement (Klimesch, 1999), Beta (18-21 Hz) – movements, Gamma (30-60 Hz) – memory and perception. In infants, these functional frequency bands entail lower frequencies. More synchronised oscillations of a certain frequency reflect greater excitability of the neuronal populations and hence more efficient neural communication. Changes in components of the ongoing oscillations in a frequency band can provide insights about cognitive functions. Components that can be extracted are the amplitude, the power and the phase of the signal at a given frequency. The difference in oscillatory power at a frequency over certain time windows in respect to a baseline time window can be calculated by frequency analyses (e.g., Fourier analysis). It can be visualised in a time-frequency spectrum, by transforming the signal from the time domain, in which the signal is depicted at each timepoint, to the frequency domain, in which phase, amplitude and power are shown for the different frequencies averaged across at time window.

Advantages of EEG. EEG is particularly suited for infant studies. The EEG system consists of electrodes that are in contact with the scalp by application on an infant-sized cap or net. It is non-invasive, portable and light weight, allowing the infants to move freely. It directly measures the neural activity at the scalp and thus has a high temporal resolution, thus typically providing samples at every few milliseconds on average, with the exact sampling rate depending on the EEG system. It is thus able to measure rapid changes in neural activation, and also not require time for a response function to build up. Minimising the signal processing time between stimulus presentation blocks is important in neurodevelopmental research, where time is limited due to behavioural tendencies of the participants. This circumstance is particularly useful in the context of real-time analysis, where the signal is being processed while the experiment is still ongoing. On the other hand, EEG provides low spatial resolution, because it only measures the signal that is present at the scalp.

2.5 The use of Neuroadaptive Bayesian Optimisation for infant EEG

The present thesis aimed to extend the Neuroadaptive Bayesian Optimisation approach to studying EEG responses in the infant brain in the context of social development. Addressing the restrictions of the traditional approach listed above, NBO is particularly suitable for the field of developmental neuroscience: First, testing multiple conditions simultaneously is useful in infant research where participants have natural behavioural tendencies (e.g., short attention span, becoming fussy, hungry and tired more quickly and more unpredictably, refusing the cap) that make the recording of a clean signal more challenging than in adult studies. Second, because the age range confining the period of infancy is relatively short, the population is much smaller than in adult research, making efficient experiments even more beneficial. Third, mapping how a certain brain metric is differentially triggered by various stimuli is particularly informative in a relatively young field in which knowledge about relevant stimuli are still scarce. Fourth, because infants not only differ in their preferences towards certain stimuli, but also in how these preferences develop in infancy, a period characterised by substantial changes on the level of brain and behaviour, NBO allows to study neurodevelopmental pathways in a unique way. Fifth, mapping how the individual's brain response varies across stimuli is particularly informative in samples with elevated likelihood for neurodevelopmental conditions, because instead of asking how an individual or subgroup differs in the response to the stimuli preferred by typical infants it asks what an individual or subgroup at elevated likelihood for atypical social development would prefer, thus providing the basis for individualised interventions and targeted support at an early phase in the developmental cascade. Finally, the requirement to set all analysis parameters before collecting the data carries particular value for the field of neurodevelopmental research which has yet to establish standardised analysis pipelines.

Together, NBO may have great utility for studying early social development on the level of individuals/subgroups. The remainder of the present chapter describes each component of the NBO loop in more depth, while Chapter 5 and 6 each present an example study applying NBO to infant EEG data.

2.6 Neuroadaptive Bayesian Optimisation with infant EEG – The experimental loop

Before the experiment, the stimuli are created and arranged along one or several dimensions. The closed experimental loop (Figure 2.1) starts with the presentation of a stimulus and recording of the EEG data. After the last trial / epoch of a stimulus presentation block, the raw EEG signal is being pre-processed and analysed to obtain the predefined target response value (e.g., in Chapter 5, the amplitude of the mean Negative central (Nc) component to the infant ERP response). This value is passed to the algorithm which maps it to the presented stimulus within the stimulus space and updates its model of the unknown brain response function across the space. Aiming at identifying the stimulus in the space that elicits the target brain response maximally, it selects the next stimulus to be presented and collected empirical data for to further inform the model. This loop continues until a user-defined stopping criterion has been reached, under which the optimum stimulus is considered as identified, or until a maximum number of runs/iterations have been reached for the case that the early stopping criterion was not reached. The time between a two stimulus presentation blocks amounts to approximately 6 seconds; ~ 1 second for the real-time analysis and ~5 seconds for the Bayesian optimisation. The exact amount of time needed depends on various factors, including the stimulus file format (longer for heavier files), size of the search space (longer for larger spaces), and the processing speed of the laptop. The 6 seconds were measured with a setup including image stimuli (.png), a 2-dimensional search space of 16 points, and a MacBook-Pro hosting stimulus presentation, real-time analysis and the BO algorithm. In the following section, the creation of the stimulus space and the and two processes comprising the loop, real-time analysis of the EEG data (section 2.8) and Bayesian Optimisation (section 2.9) will be described in detail.

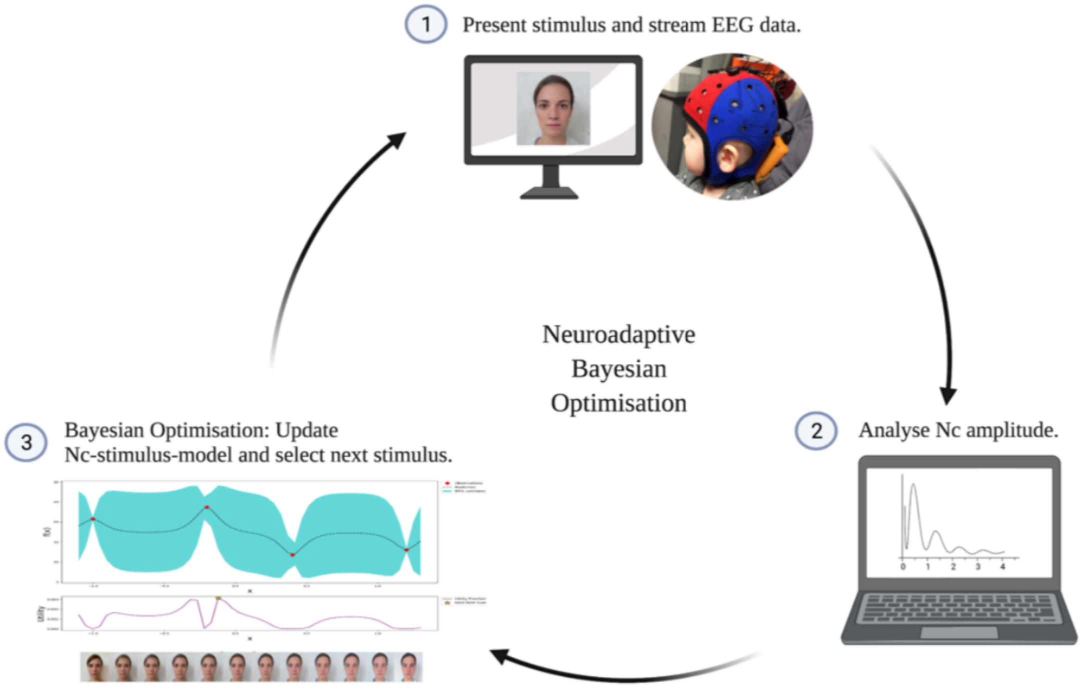


Figure 2.1. Example loop of Neuroadaptive Bayesian Optimisation with infant EEG data

2.7 The stimulus space

The Neuroadaptive Bayesian Optimisation iteratively maps sampled empirical values to the respective stimulus presented from a prepared stimulus space, to iteratively build up a predicted target brain response across the space. In the stimulus space, the stimuli have to be meaningfully arranged across one or more stimulus dimensions. The stimuli can either be ordered by the intrinsic stimulus dimensions (e.g., familiarity), or by predicted responses based on accumulated previous research findings of the same or a different but proposedly related response metric (Lorenz et al., 2017).

Stimuli can then be ordered linearly along a dimension, varying from one extreme to the other, or with peak characteristics of the dimension in the middle of the space and minimum characteristics at the extremes, or the other way around. The meaningful arrangement of the order of stimulus points along a dimension allows the BO algorithm to more easily learn the unknown brain response function. The number of possible stimuli per dimension is theoretically endless, and the number of dimensions, too.

As an example of a stimulus space, in the proof-of-principle NBO study presented in Chapter 4, a range of faces varying in their degree of similarity to the parent's face were ordered along a linear continuum ranging from very familiar to very unfamiliar.

2.8 Real-time analysis of infant EEG data

Real-time analysis of neurophysiological data in the context of NBO comes with several additional requirements compared to offline analysis. First, the data has to be streamed rather than merely recorded, in order to be available for analysis after the end of each stimulus presentation block. Second, all steps of the analysis pipeline have to be pre-scripted in order to run after each block. This includes the choice of a target brain metric to optimise as well as choices regarding the various parameters of the pre-processing steps, which have to be made before starting data collection. These decisions can be based on a combination of established choices commonly used the field and/or offline exploration of pilot data collected with the specific paradigm and set-up.

Of note, the following sections relate to choices of EEG real-time processing parameters in the context of the NBO experiments presented in Chapter 5 and 6; the parameter choice for the respective study can be found directly in the chapters, while here the rationale behind selecting these parameters is described. The pre-processing of the EEG data used in the traditional analyses presented in Chapter 3 and 4 used slightly different parameters, which are specified in the respective chapter, too.

2.8.1 Choice of EEG system, data streaming and synchronisation.

2.8.1.1 Data streaming

There are multiple components in the real-time loop and changing one of the components can affect the functioning of the others:

- EEG amplifier – here: *Enobio box by Neuroelectrics*
- EEG acquisition software – here: *Neuroelectrics Instrument Controller (NIC2)*
- Streaming software – here: *Lab Streaming Layer, LSL*
- External software hosting stimulus presentation and EEG data analysis – here: *MATLAB*
- External software hosting Bayesian Optimisation – here: *Python*

MATLAB presents a stimulus on the participant's screen and sends a stimulus marker to the acquisition software via LSL. The amplifier receives data from the electrodes on the scalp and sends them to the acquisition software via cable or network. The acquisition software receives the EEG data from the amplifier and the marker from MATLAB. Due to the wireless nature of the Enobio system, there might be some latency in the data arriving in the acquisition software from the amplifier, relative to the clock of the acquisition laptop. The acquisition software compensates this latency and aligns the two clocks. When the acquisition software receives markers from other applications, a third clock is to be synchronised, which is the clock from the host sending the markers. In order to gather these markers, Lab Streaming Layer (LSL) can be used, which is able to synchronise data with an accuracy of 1 ms. LSL assigns a common time stamp to EEG data and marker. Finally, MATLAB pulls the synchronised EEG data and marker chunks from LSL.

The marker-aligned EEG data is analysed in MATLAB and an output value is written to an external text file saved on the local network. The external software hosting the BO algorithm (e.g., Python) that monitors the text file reads in the value and runs the Bayesian Optimisation which maps the value onto the stimulus space to update the predicted function. Taking into account all sampled responses, it selects the next stimulus to present and copies the respective stimulus file (e.g., an image file) to a dedicated stimulus folder. This stimulus folder is being monitored by MATLAB, which reads in the stimulus file from the folder and presents it on the participant screen. This loop is repeated until the early stopping criterion or a pre-defined maximum of blocks is reached.

2.8.1.2 EEG system

The Neuroelectrics ENOBIO EEG 8-channel system was chosen due to its being wireless and lightweight and because it is compatible with Lab Streaming Layer.

The Neuroelectrics NE Enobio device is a wireless EEG amplifier that connects to its acquisition software via USB cable, WIFI, or Bluetooth. The type of paradigm influences the choice of connection type. Due to its being wireless, it allows for some freedom of movement, which is convenient when recording EEG data from infants. In the real-time context, this is even more relevant. In the stimulus presentation breaks during which the data was being analysed and the next stimulus selected (~6 seconds), infants were able to move freely. In the case of live social interaction, this advantage is particularly highlighted as the infant may move around more. Furthermore, the NE acquisition software, Neuroelectrics Instrument Controller (NIC2), is compatible with LSL for data streaming, which is necessary for the analysing the data in real time.

The Enobio system uses two referencing electrodes that serve as the “ground”: the “Common mode sense (CMS)”, and the “driven right leg (DRL)”. All signal is measured in reference to CMS. The DRL brings the measured potential as close as possible towards “zero”. Provided that there is good contact of the reference electrodes to the skin, they reduce electrical noise and the drift of the signal.

2.8.1.3 Time-stamp synchronisation during EEG data streaming

The synchronisation between EEG data and marker time stamp is not perfect. Rather, the time at which the marker was sent from MATLAB might not always match with the time stamp assigned to the EEG data. This is the case, for example, if the stimulus presentation is slightly delayed due to factors such as screen refresh rate, or due to changing network properties, especially when using wireless connections. The mismatch can go in both directions, that is the markers can be delayed relative to the EEG signal, or they can be early relative to the EEG signal. A delay of the markers relative to the EEG signal arises when the markers sent from the MATLAB script take longer to arrive in the NIC acquisition software take longer than the time it takes for the EEG data to be transmitted from the amplifier to the acquisition computer. Likewise, the marker time stamps are earlier than the EEG signal when the EEG signal takes longer to arrive in NIC than the markers. The gap might further increase in the course of the session, that is with the time that has passed since the initial synchronisation in NIC; for example, it might be influenced by the time the NIC software has been running.

Therefore, before starting the present studies, the potential delay of the true stimulus onset relative to the ENOBIO marker was calculated. To do this, a light sensor was used as marker in addition to MATLAB-sent markers and calculated the difference in their respective time stamps. The light sensor signal is perfectly accurate since the light patch is presented exactly then when the stimulus is being presented, and both light signal and brain signal are jointly measured and transmitted to NIC. Therefore, the marker delay with light sensor is zero. However, with infants, the light sensor cannot be used as a marker, because the electrode measuring it is (just as the other electrodes) influenced by the noise measured in the ground electrode (because all electrodes are measured in relation to this). With infants, due to enhanced movement, there is more noise in the ground electrode, and therefore in the LS electrode. Hence, the light sensor would be less likely to pick up the light signal, as a function of the amount of movement. Therefore, in pilot sessions, the light sensor has proved to be a reliable marker with infants, and the MATLAB markers were used in the NBO studies instead.

The mismatch between MATLAB markers and light sensor was computed under different conditions and including different task durations: with restarting NIC just before acquisition versus having NIC running for 25 minutes before starting the acquisition; and with WIFI versus USB cable connection to send data from amplifier to NIC. When using WIFI, the variance of the mismatch was greater than when using USB cable connection, while the mean mismatch was roughly the same across connection modes (Figure 2.2). Because it is difficult to correct for a varying delay in the experiment, the USB cable was chosen for the ERP paradigm where exact timing is critical. The delay was corrected for by adding the mean duration of the delay (35 ms) to each marker time stamp.

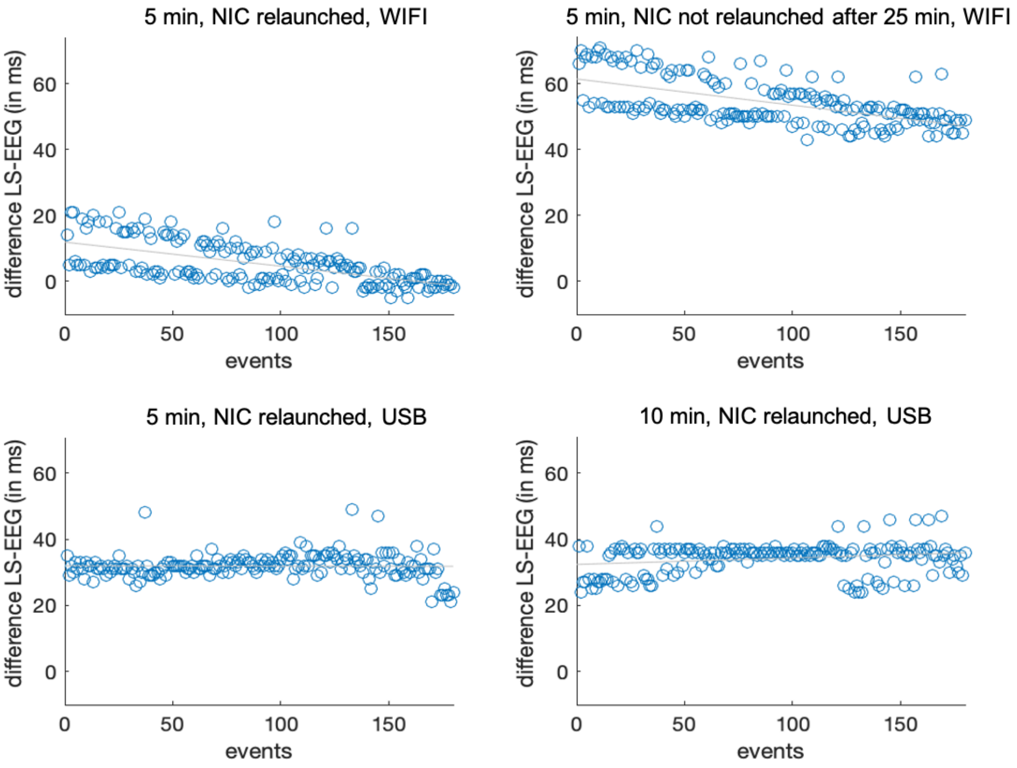


Figure 2.2 Differences between the light sensor time stamp and MATLAB marker time stamp across time, under different technical conditions, incl. 5 min versus 10 min experiment duration, NIC2= Neuroelectrics Instrument Controller acquisition software

2.8.2 The target response

The pre-defined brain response is being selected based on previous literature. Depending on the research question, this can be any brain metric measurable with the chosen system that is expected to be relevant for the given research question and to be sensitive to the variation in the stimulus space.

If there are multiple options of possible brain response features, and strong theoretical predictions of how the response function should map on the search space on an averaged, group level, it might be helpful to analyse pilot data from a group of infants by computing the different metrics and comparing them in how they differentially map onto the stimulus space. It is thereby crucial that the target metric reliably captures the process under study, that is that similar values are computed upon repeated sampling at the same area in the space.

2.8.3 Defining the amount of data needed per iteration/block

2.8.3.1 Number of trials/epochs

In an event-related design, such as an ERP paradigm, this is the number of trials of brief presentations needed in order to obtain a reliable signal. In a block design, such as when studying EEG oscillations, this is the number of seconds or epochs for which the continuous stimulation is ongoing. Previous research showed that more trials are required if there is noise in the data and if the ERP component is relatively small (Picton et al., 2000), and that the number of trials increases reliability of ERPs in adults (Huffmeijer et al., 2014). Especially in infants, where EEG data shows more artifacts, and hence ERPs show greater within-subject variance (De Haan, 2007), a higher number of trials likely improves reliability. Typically, between 10 and 15 trials are required in visual ERP paradigms (Hoehl & Wahl, 2012).

However, the decision of the right number of trials in infant EEG studies is not trivial. On the one hand, too few trials might result in the EEG metric not being robust enough in order for it to be reliable, hence preventing the BO from convergence. On the other hand, an increasing number of trials and hence increasing duration of seeing the same stimulus repeatedly, infants' interest will decrease, and the risk of not reaching enough clean trials increase. In an NBO experiment, this means not being able to continue due to bad data quality and having to stop before the paradigm is completed. Therefore, the amount of data needed per block should be identified during the pilot phase of the study. The test-retest reliability of the individual infant Nc between two visits has been shown to be lower than of face-sensitive ERP components in adults but still moderate with 10 trials (Munsters, van Ravenswaaij, van den Boomen & Kemner, 2019). Considering that during an NBO experiment reliability is needed not across two visits taking place on different days, but within the same session, just minutes apart, test-retest reliability can be assumed to be even higher here. One way to evaluate how many trials or seconds will be enough, is to offline re-analyse raw data acquired in an NBO session. Notably, NBO to some degree integrates trials across blocks. After each block, it updates its prediction about the optimum, taking into account all values sampled up to that point. In other words, the prediction about the optimum is not based on only the trials of one block, but of all trials across previous blocks, to an extent compensating the relatively small number of trials within each block.

2.8.3.2 Number of blocks

While the researcher can set a maximum number of blocks after which the experiment should be terminated, the actual number of blocks of the experiment for a given child depends on the timepoint of convergence of the BO, which in turn depends on the reliability of the brain signal obtained. Defining a maximum number of blocks, after which the BO, if not already converged, will stop sampling and the experiment is finished, is sensible to constrain the maximum length of the experiment. With infants, the maximum attention span during screen-based experiments has been shown to be about 20 minutes. Therefore, a maximum number of blocks equivalent in duration to 20 minutes, after which the experiment will end, may be chose. Of note, the BO builds its model on the basis of an initial model obtained during the burn-in phase when pre-defined stimuli from the space are presented. Therefore, the number of burn-in blocks plus a few iterations are the minimum number of blocks needed for an NBO experiment. Depending on the stopping criterion for identification of the optimum (here: sampling of the same stimulus in three consecutive iterations), this additional fixed number of blocks needs to be considered.

2.8.3.3 Number of participants

The sample size depends on the exact research question. In order to only obtain individual preferences of infants, no minimum sample size is needed as no group comparison is done. However, if the individual results should be used for further analysis to answer questions on the (sub-)group level, then the sample size should be determined a priori, just as is done in traditional experiments (e.g., by computing a power analysis).

2.8.4 Choice of pre-processing parameters

2.8.4.1 Filtering

The EEG signal includes low frequency noise and high frequency noise. Low frequency noise can be caused by head movement and scalp persipiration and manifests as drift in the signal. High frequency noise can be caused by face muscle contractions or electrical noise in the room, and manifests as very small and fast ups-and-downs in the signal. This noise can be reduced by filtering the signal from noisy frequencies, that is excluding frequencies below and above a certain frequency range of interest. A high-pass filter passes signals with a higher frequency than the cut-off and attenuates lower-frequency signals, and a low-pass filter passes signals with lower frequencies than the cut-off and attenuates higher frequencies.

In the current NBO studies (Chapter 5 and 6), low-pass filters were applied to filter out high-frequency noise and 50 Hz line noise coming from the EEG apparatus (Chapter 5: 20 Hz; Chapter 6: 35 Hz). Different high-pass filters were considered, including 0.1 Hz and 1 Hz. The 1 Hz option seemed preferable, since it reduced drifts observed in the signal that diminished the Nc negative amplitude. On the other hand, 0.1 Hz was considered as an option because it is more commonly used in the literature (e.g., De Haan & Nelson, 1997; De Haan & Nelson, 1999; Webb et al., 2011) and further does allow more data to pass, hence affecting the signal less. I compared the shape of 44 ERPs produced with a high-pass filter of either 0.1 Hz or 1 Hz. I also included detrending versus not detrending in the comparison in the search for a way to reduce drifting. Nc shapes were judged by their characteristic negative deflection starting after 250 ms and returning to crossing zero before 800 ms. These Nc shapes were best resembled by ERPs produced using a high-pass filter of 0.1 Hz and detrending of the signal. Detrending reduced drifts in the signal. The more inclusive high-pass filter threshold of 0.1 Hz allowed us to ensure not to exclude relevant data by cropping the ERP amplitude (see *A.2.1* for plots).

2.8.4.2 Artifact rejection

Automatic artifact rejection requires careful choice of the rejection criteria. In the current work, this choice was supported both by common procedures and extensive piloting, in order to reveal which parameters were able to produce a reliable signal, that is whether in similar areas of the stimulus space similar values are being sampled on the individual level. This test-retest-reliability is a requirement of the success of the BO: only reliability in the values allows the BO to reduce uncertainty and identify the point in the search space robustly eliciting the strongest response.

Of note, different artifact rejection thresholds might be suitable for different participants, based on the individual characteristics of the general brain response. For example, some infants show naturally larger ERPs than others, and a common threshold for artifact rejection would result either, if a more lenient threshold is chosen, in including too much noise in the signal for infants with naturally smaller ERPs, or, if a more conservative threshold is chosen, in discarding valuable data from infants with naturally larger ERPs. Therefore, it may be helpful to assess an individual infant's general ERP size at the beginning of the study and adjust artifact rejection thresholds accordingly. Because the interest lies in the relative difference in ERPs to different stimuli within the same individual, even if later compared with other individuals in a group analysis, under the NBO framework this is not problematic.

2.8.4.3 Dealing with eye blinks

One challenge faced in real-time experiments is the question of how to deal with eye blinks. In classic studies analysing the data offline after completion of the study, researchers can manually identify per video footage of the participant's behaviour moments of blinks and exclude the respective trials from the analysis. In real-time experiments, this is not possible, because there is not enough time to review the video footage and manually exclude trials.

A helpful tool for at least partially *preventing blink artifacts* in the data is gaze-contingent stimulus presentation. Through video-monitoring, the researcher can observe the participant's looking behaviour, and only present the stimulus once the participant is looking towards the screen with open eyes. While this is mainly aimed towards excluding trials of looking away, it can also help to prevent from including trials when the child was blinking. However, this only partially solves the problem since it does not prevent from including trials in which a blink occurred *after* trial onset.

If blinks cannot be fully prevented, a method is needed to deal with the blink artifacts present in the data. One way to detect blinks automatically are *threshold-based artifact detection methods*. Trials would then be identified as blink trials if the signal resembles the typical topographical characteristics observed in blink artifacts – a brief, spiking rise and subsequent fall in amplitude. Threshold based blink detection is difficult with infant EEG data because infants' EEG stream is of generally higher amplitude (Bell & Cuevas, 2012) and therefore the magnitude of the difference between the amplitude of the blink signal and the brain signal is smaller, making detection by amplitude threshold more difficult. Further, individuals differ in their blink duration (e.g., Benedetto et al., 2011). Hence, a common latency threshold might not fit all individuals, too. The individual blink duration would first have to be established at the beginning of the experiment; waiting for a blink to occur and via video monitoring and key press to manually identify the timepoint of the onset of the blink. With infants, given their short attention span, this step would risk sacrificing valuable time that might later be missing for completion of the paradigm. With infants, this problem is further enhanced, because their blinks are less stereotypical than in adults and hence even more difficult to identify in the EEG data (e.g., Fujioka, Mourad, He & Trainor, 2011). There are other, more principle problems to threshold-based artifact detection – for example, if a blink happens across two trials, with the peak happening in one trial and the amplitude not reaching the threshold in the other trial, this other trial would falsely not be detected as blink-contaminated (Klein & Skrandies, 2013).

As another way to detect blink artifacts automatically is *Independent Component Analysis (ICA)*, which has been suggested to be more favourable than threshold-based methods because it is not limited to a single feature and can work on the basis of one individual's data. However, is on the one hand difficult with infant EEG data, and on the other hand impossible in a real-time context. Infants show high intra-subject variability, more than adult, due to more frequent and more abrupt movements, creating artifacts in the data (e.g., Hoehl, Wahl, Michel & Striano, 2012). Hence, ICA with infant EEG data needs a higher amount of input data than with adult EEG. In a real-time context each block is being analysed separately, providing for analysis only a small amount of data is available for analysis, making reliable ICA for blink detection impossible, especially in infant data, in which by tendency more data is needed due to higher variability in the data. In principle, ICA would become more possible later on in the course of the experiment, with data across blocks accumulating, but not in the first part of the experiment, and then still it remains to be tested whether the amount of data suffices to reliably extract blink components from the infant's EEG stream.

Finally, a further way to detect blinks is using an *electrooculogram (EOG)*, that can be used in combination with EEG. The trials marked by the EOG as blink-contaminated could then be discarded from the EEG data. However, with infants I want to keep the recording system as non-intrusive as possible and the NE Enobio system does not include an EOG electrode.
Overall, none of the common possibilities for artifact-detection in EEG data seems to work with infants in a real-time context.

Given the lack of effective and reliable automatic blink detection algorithms which could be implemented in a real-time pre-processing pipeline, it is worth considering not controlling for blinks at all under evaluation of the potential effects on the output measure. For example, for the real-time paradigm investigating infants' Nc mean amplitude to parent's versus stranger's face (Chapter 5), I assessed the possible effects blinks could have on the Nc mean amplitude.

First, blinks are generally relatively rare in infants. For example, 7-months old infants blinked 3.5x/min while looking at their mother's face and 6x/min while looking at a stranger's face (Tummeltshammer, Feldman & Amso, 2019), that is a maximum blink rate of every 10-17 seconds. One block in the current paradigm takes 18 seconds, if the child is not looking away from the screen. Therefore, I expected blinks in only 1-2 trials maximum out of the 12 total trials per image, depending on whether the face resembles more the mother or the stranger.

Second, it is worth using pilot data to quantify the effect of blinks on the outcome measure. For example, I used pilot data to estimate the effect of blinks on the Nc mean amplitude in the ERP paradigm (Chapter 5). Video footage of 564 trials collected from 5 infants showed that blinks were very rare in this paradigm with that particular age group. In total, blinks were observed in 2.3 % of the trials. The Nc values from the blocks containing blinks were subjected to a Welch two sample t-test, revealing that mean Nc values obtained per block did not differ depending on whether or not the blink trials were included in calculating the Nc mean amplitude of the block (t = 0.11, df = 13.98, p = .9; mean of block average when including blink trials: -25.43; excluding blink trials: -25.93). The reason for that was that these trials often did not survive automatic artifact rejection and were hence anyway not included in the ERP. In sum, I assumed that blinks would not affect the measure of interest, since blinks are expected to be rare and the blinks observed in the pilot data did not affect the mean Nc amplitude on the block level.

Finally, it is helpful to evaluate what, as worst case, the effect of frequent blinking could have for the interpretation of the outcome. For example, applied tothe present paradigm (Chapter 5), since blinks are characterised by high amplitude values in the EEG signal and given that infants blink more while looking at a stranger's face then looking at mother's face (Tummeltshammer, Feldman & Amso, 2019), blinks should inflate the Nc signal for stranger versus mum. If the Nc difference would still be in the expected direction, with stronger Nc for mum versus stranger, blinks would if anything support the success of the approach.

It is subject to future research to find out whether manual blink indication, by the experimenter via video monitoring and key press, could provide a reliable way of eliminating blink-contaminated trials. In this scenario, one processing step would be concerned with evaluating whether a dedicated key was pressed in the course of each single trial, and if yes, exclude the respective trials from the analysis. It remains to be tested in the specific paradigm (with the specific stimulus presentation time and specific researcher) whether reaction times are short enough to allow for reliably identifying blinks, or whether reaction times would be too long resulting in good trials being accidentally be excluded and contaminated blinks be missed. It is also subject to future piloting to find out whether EOG, in combination with EEG, would be tolerated by the infant and be able to reliably remove infant eye blinks from the data.

Together, given the lack of suitable blink artifact detection methods for an infant real-time EEG paradigm, the potential effects of blinks should be carefully evaluated before the start of the study before a decision is made for or against automatic blink control. For example, eye blinks were not controlled for in the real-time ERP study beyond gaze-contingent stimulus presentation (Chapter 5), because their frequency and effect on the mean Nc amplitude was expected to be minimal, and the use of automatic blink rejection steps would have introduced new problems.

Of note, a recently published study reports the development of an algorithm that is able to identify artifacts in ERP data in a real-time neuroadaptive approach (Ouyang et al., 2022). Future NBO infant studies should test this method with infant EEG data.

2.8.4.4 Re-referencing the signal

Electrical signal coming from a dipole must go into a ground. EEG amplifiers create a virtual ground where the signal from the active electrodes arrive. Hence, all signal from the active electrodes are subtracted by the ground signal. However, the ground signal picks up electrical noise from the amplifier that does not reach the other scalp electrodes, and thus the difference between the active signal and the ground signal is also affected by this noise. To cancel out this ground-related noise, a reference electrode is chosen (the signal of that electrode is traveling to the ground, too). By subtracting the reference electrode from the active electrodes, that ground noise can be eliminated (because [A-G]-[R-G] = A-R). Some amplifiers ("differential amplifiers") do the referencing step online, during recording. In that case, the output of the amplifier is the difference between the signal in the active channels and the signal in the online reference. Other amplifiers (including ENOBIO by Neuroelectrics, used in the present NBO studies) lack an online referencing step. Instead, their output is the raw ground-subtracted signal. The referencing step can be done in software after recording. Neuroelectrics recommends doing differential referencing in order to improve the signal quality.

To decide whether re-referencing should be applied in the real-time experiment, pilot data from the specific paradigm can be used to calculate the target output value including re-referencing versus not including re-referencing and assess the reliability of the sampled values in both scenarios. The calculations can include different re-referencing options, to see which channel combination produces the most reliable results. Further, it can be helpful to investigate how different re-referencing options affect how the target response maps on the search space, aiming to achieve a meaningful variation of the target response across the space. For example, if the prediction is a stronger brain response towards one versus the other extreme of a 1-dimensional stimulus space, then the difference in target response to these two extracted points averaged across pilot participants can be re-calculated using different re-referencing options, to see which produces the most differential result in line with the hypothesis.

2.8.4.5 Real-time data quality control

An important aspect to consider prior to the study is how to make sure that the output value is reliable before it is being passed to the BO which will integrate it in its surrogate model. Unreliable output values mapped onto the search space will result in ill surrogate models with large uncertainty, inhibiting convergence to an optimum. Ensuring reliability in the output measure in relation to the search space firstly has to be ensured by a strong background of literature showing that the output measure reflects the cognitive process under study. Further, a minimum number of trials / epochs should be involved in the average of the final output metric. To this end, a certain number of trials is being presented to the participant which already assumes a certain percentage of data loss. In addition to these considerations, it is useful to implement a real-time control of the data quality.

In traditional studies, the data quality is being examined after the study and ERPs are discarded if not at least a certain number of valid trials is involved in the average. In a real-time scenario, this check has to be carried out during the experiment, after each block of data collection. If after a given block the number of valid trials is below a certain threshold (e.g., 10 trials), the data should be discarded, and data collection repeated for that block. In the ERP study presented in Chapter 5, the script calculated after each block the percentage of how many trials across channels survived artifact rejection and went into the average of the ERP. If this percentage was lower than the equivalent of 10 trials aimed for, the data of that block was discarded, and stimulus presentation and data collection repeated. Of note, in order to make use of as much valid data as possible, the 10 "good" trials aimed for could be recorded from different channels of interest. In other words, if one channel of interest provided 10 valid trials, the threshold was met as well as if the six channels of interest each provided a few valid trials.

In addition to knowing how many trials survived artifact rejection in a given block, it is informative for the researcher to know for how many trials in each of the channels of interest this was the case, in order to undertake steps to improve signal quality in the respective electrode. To this end, in the real-time studies presented in this thesis included a feedback to the researcher after each block, plotting the number of valid trials per channel of interest.

2.9 Bayesian Optimisation

The Neuroadaptive Bayesian Optimisation approach presents an individualized alternative to the standard experimental approach. It powerfully samples empirical points to find extrema of an unknown function $f(x)=y$, such as the maximum ERP amplitude response within a given stimulus space *x*. In a closed-loop manner, it iteratively selects stimuli from the space to present and collect an empirical response value for, then maps the empirical response onto that point in the space and updates the model of the unknown function to the sampled values.

An initially defined acquisition function guides which point to sample next, based on the previously sampled values and the uncertainty across the space. Because it balances maximising the predicted function, that is *exploiting* the space where previously sampled values were highest, and minimizing uncertainty, that is *exploring* the space where the uncertainty of the predicted function is highest, it is able to find function extrema after a few iterations, after only presenting a subset of stimuli from the space to the participant, which is particularly useful in contexts where sampling at each point in the stimulus space is very costly such as in infant studies.

The BO algorithm consists of two components: the "surrogate model" and the "acquisition function". Over the course of the experiment, the algorithm fits a statistical model ("surrogate model") to the empirical values, aiming to model the true, unknown response function across the entire space. After each iteration, the surrogate passes the point where it predicts the highest value and the model's standard deviation to the acquisition function.

The "acquisition function" determines the point in the stimulus space that is going to be sampled in the next iteration based on the point in the space where the highest value is expected, and where the uncertainty about the true underlying value is greatest. Which of these aspects the acquisition prioritises depends the user-defined aim of rather mapping out the entire space (*exploration*), *or* rapidly identifying the maximum while ignoring the rest of the space (*exploitation*).

NBO requires not only pre-selection of signal processing parameters prior to the experiment, but also parameters for the optimisation algorithm itself. Piloting the pipeline helps to identify the optimal parameters for the respective study. The following paragraphs will describe which parameters for each BO components were chosen for the studies presented in this thesis. For the studies presented in this thesis, the Bayesian optimisation algorithm is hosted in Python.

2.9.1 Surrogate model / Objective function

The surrogate model, or the objective function, is a statistical model that is fit to the empirical values obtained from iteratively sampling points in the stimulus space. The model aims to represent the true, underlying, but unknown response function of the individual across the stimulus space. In the current thesis, a Gaussian Process Regression model $GP(x) = p(x|y)$ was used as surrogate model, with a stationary Mátern kernel and a smoothness parameter of $\nu = 2.5$. A white noise term was added to estimate the noise level in the data (for details see Da Costa et al., 2021, p. 11).

2.9.2 Acquisition function

The acquisition function, or utility function ($u(x)=\max(0,f' - f(x))$, determines the point in the stimulus space that is going to be sampled in the following iteration (for a detailed description, see also Da Costa et al., 2021). There are different options of utility functions that can be chosen for a paradigm depending on the respective research question. One option is the *Expected Improvement (EI)* acquisition function that was used in the present studies (Frazier, 2018). The Expected Improvement acquisition function is the expected utility as a function of x ($E[I(x)]=E[\max(f(x^{*}) - y, 0)]$, Da Costa et al., 2021, p. 11). It selects for sampling the point with the highest expected improvement (i.e., the maximal expected utility). It has two components that play a role in evaluating which point is evaluated to have greatest utility in being sampled next. The first component regards maximising/minimising the mean function $\mu(x)$, that is where in the space the highest/lowest value is expected. The second component regards minimising the variance, that is where the uncertainty about the true underlying value is greatest. That is, it samples f at the point that it expects to improve upon f' the most.

By considering both components, the Expected Improvement utility function offers a trade-off between exploitation (evaluating at points with high mean) and exploration (evaluating at points with high uncertainty). The user can pre-define which of these aspects is prioritised depending on either aiming to rapidly identify the point where the highest value is expected or extensively mapping out the entire space.

With the the surrogate function following a normal distribution, EI can be expressed as (see Da Costa et al., 2021, p. 11):

$$E[I(x)] = (\ \mu(x) - f(x^{*}))\ \Phi(z) +\ \sigma(z)\ \phi(z), where\ z = \frac{\mu(x) - f(x* + \xi)}{\sigma(x)}$$

whereby f(x*) is the predicted value at point x in the space, ϕ is the standard normal density and Φ is the standard normal distribution function, reflecting features of the normal distribution of the surrogate function. ξ is the hyperparameter determining the balance between exploration over exploitation. A higher ξ results in more exploratory sampling behaviour, that is prioritising sampling where uncertainty is highest as opposed to where maximum values are expected, aiming to minimise the standard deviation of the prediction. A lower ξ results in more exploitative sampling behaviour, that is prioritising sampling where previously highest values were found; when ξ is 0, uncertainty of the prediction is not taken into account at all.

With the Expected Improvement acquisition function, the algorithm starts sampling in an exploration mode by sampling various points across the stimulus space. Based on the user-defined aim of either mapping out the entire space or rapidly identifying the maximum while ignoring the rest of the space, the algorithm stays in the exploration mode until the end of the experiment, or after some initial iterations soon switches to an exploitation mode by sampling closely around the stimulus point that is expected to consistently elicit the highest value. Because of its ability to switch from exploration and exploitation mode, it can identify the function extrema after a limited number of iterations. This is particularly helpful in infant studies where each iteration of sampling means trying the infants' interest the presented stimulation and getting closer to the end of their short attention span.

Figure 2.3 is taken from the pre-print by Da Costa et al. (2021). Figure 2.3 A) represents an example of a surrogate model of one participant after four sampling iterations. The stars indicate where the utility function would sample next, depending on its hyperparameter ξ (here: 0.1, 0.5 or 5.0), illustrating how a larger hyperparameter ξ results in the acquisition function prioritising exploring uncertain parts of the space, while a lower hyperparameter ξ results in the acquisition function prioritising sampling where a maximising or minimising of the response (depending on the aim of the respective optimisation) can be expected. Figure 2.3 B) represents three versions of utility functions depending on the hyperparameter ξ in the acquisition function.

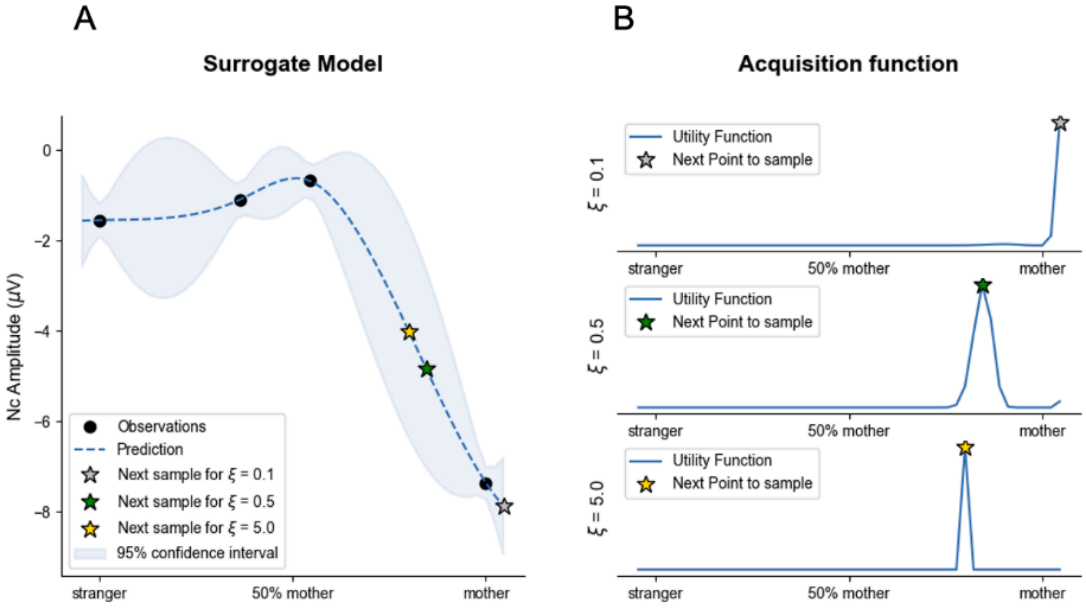


Figure 2.3. A) A participant's surrogate model after four sampling iterations, with stars indicating the next point to sample, based on different hyperparameters ξ in the acquisition function. The x-axis represents the stimulus space, the y-axis the neural target. Dots represent empirical samples taken; the dashed line represents the predicted model of the neural target across the space; the shaded area represents the uncertainty of the prediction. *B)* The utility function indicating where to sample next based on hyperparameter ξ in the acquisition function. (Da Costa et al. 2019)

Of note, the exact hyperparameter ξ value that is useful for a particular balance between exploration and exploitation depends on the size of the stimulus space and has to be identified by piloting, while considering the respective conditions and aims of the study. For example, because the duration of an infant EEG experiment can be limited by partly unpredictable factors related to behavioural tendencies of the participant, keeping the number of BO iterations low increases the chance of identifying the optimum stimulus before the end of the session. On the other hand, fewer iterations arguably mean a less robust surrogate model with a possibly less robust prediction of the optimum. In this trade-off, priorities need to be set by the researcher with regard to the aims in the respective study. It is beyond question though that high reliability of the target metric increases robustness of the prediction, in turn reducing the number of iterations needed.

2.9.3 Burn-in phase

At the start of an NBO experiment, a small number of *predefined* stimuli is sampled in order to provide the acquisition function with an initial surrogate model to start the optimisation process with. This initial phase of presenting predefined stimuli ("burn-ins") is called "burn-in" phase. The number of burn-ins and which points in the space they represent is defined by the researcher based on the respective paradigm. For the present studies I used four burn-ins, respectively, which included the extremes of the respective stimulus space.

2.9.4 Identification of the optimum

The acquisition function will determine points to sample until a user-defined condition is reached under which the optimal stimulus is considered identified, or a user-defined maximum number of tolerated iterations reached before the optimum could be identified.

2.9.4.1. Early stopping criterion

The researcher has to define a *stopping criterion* determining the condition under which the BO should consider the optimum as identified and stop further sampling. This user-defined condition should reflect that the BO algorithm is not gaining any new information. One way to determine that the BO is not gaining new information is when it keeps sampling the same point, because the acquisition function continues to determine this point as the best next point to be sampled, based on its predicted value amplitude and uncertainty. Other ways to see that the BO is not gaining new information could be, for example, if the decrease in standard deviation per iteration has fallen below a certain threshold, or if the value at one point is higher as the others by a certain percentage.

In the NBO studies presented in this thesis, as stopping criterion was defined the consecutive sampling of the same stimulus point for three times (as in Lorenz et al., 2018). Of note, while sampling the optimal stimulus beyond three times might theoretically increase the reliability of the measure, this would introduce the problem of habituation of the brain response to the stimulus and a decline in the strength of the signal measured (Nordt, Hoehl & Weigelt, 2016; Webb et al., 2010).

2.9.4.2 Maximum number of tolerated iterations

Further, it is useful to also pre-define a maximum number of iterations for which the BO should try to optimise, in order to be able to exit the experimental loop after a user-defined amount of time. While with adults, this can be a fair number of iterations, limited only by factors such as monetary compensation for study time, with infants, this depends on various unpredictable factors, but primarily the duration of their attention span, in research studies on average spanning up to approximately 20 minutes. At the end of the paradigm, whether or not convergence was met, the algorithm will output the position on the search space estimated to produce the strongest target response for the respective individual. These space coordinates of the preferred stimulus can be used in further analyses. The optima derived this way are still informative, because the algorithm updates the predicted model of the underlying function after each iteration of sampling, taking into account all samples collected of an infant up to that point. Hence, after the final block, that was the optimum the algorithm predicted based on the collected data. The fact that no stopping criterion was reached in these cases is primarily relevant for the course of the experiment itself – in infants, for example, given their relatively short attention span, it is helpful to only sample as much data as needed. The fact that no stopping criterion was reached also implies that the uncertainty of the prediction is to some degree higher than if the stopping criterion was reached. In order to account for this fact, an extra-step could be to define a threshold of certainty of the predicted function which has to be met in order to be included for further analyses.

In the present NBO studies, the maximum number of blocks was set to 15 (i.e., 4 burn-ins + 11 optimisation iterations), equivalent to approximately 20 minutes, which I considered as the infants' maximum attention span.

2.9.5 Summary

NBO is an alternative experimental approach using a closed-loop design to model an unknown function and rapidly identify extrema in this model while only presenting a subset of stimuli. The acquisition function of the optimisation algorithm guides which stimuli to present, based on features of the surrogate model. The loop will stop once a user-defined stopping criterion has been met. Because of its ability to switch from exploration and exploitation mode, the BO is able to identify the function extrema after a limited number of iterations, which allows to identify individual maxima in neural responses of infants.

Parameters the researcher has to define prior to the experiment are whether to maximise or minimise the output value, the number and space positions of the burn-in stimuli, the balance between exploration and exploitation *(ξ)*, the stopping criterion under which the loop should be exited considering the optimum as identified, the maximum number of iterations tolerated before the loop should be exited in case the early stopping criterion has not been reached.

In the present thesis, the Bayesian Optimisation algorithm was used for adaptive sampling of neurophysiological data to identify the stimuli maximally triggering responses in an infant's social brain network. In a first proof-of-principle study, the aim was to validate whether NBO is able to pick up meaningful differences in the infant EEG target brain response (Chapter 5), while in a second experiment, the method was applied to a naturalistic context with a 2-dimensional stimulus space.

2.10 Conclusion

This chapter introduced NBO as a new experimental approach that is able to test multiple predictions at the same time by mapping responses across a wide stimulus space, without the need of presenting each single stimulus. It can thus be used to efficiently study the meaning of brain functions and variability in the stimuli that optimally trigger brain responses in different subgroups and individuals. Due to its efficacy, and combined with the method of EEG, it is particularly valuable for application in infant studies.

NBO can be understood to complement the traditional experimental approach. While the traditional approach studies how on the group level brain responses are modulated by few preselected stimuli, NBO asks which of a range of possible stimuli is optimal for a subgroup or an individual to elicit a relevant target brain state, and the two approaches can mutually inform each other by their different perspectives and relative strengths.

Of note, this methods chapter focussed on extending the novel experimental approach Neuroadaptive Bayesian Optimisation to infant neurophysiological data. Studies applying this new method are presented later in this thesis (Chapters 5 and 6), while Chapters 3 and 4 use the traditional experimental approach of averaging responses across the sample.

CHAPTER 3: THETA POWER RESPONSES DURING LIVE SOCIAL EXPERIENCES IN INFANTS WITH TYPICAL AND ELEVATED LIKELIHOOD FOR AUTISM AT 14 MONTHS

3.1 Introduction

Difficulties in social communication and interaction are common characteristics of autism. While a diagnosis can be given in early childhood, alterations may have emerged before they become visible through enhanced social demands. In order to understand the path to behavioural symptoms, it is important to identify differences on the level of neural and cognitive processing early on. Naturalistic social contexts as opposed to pre-recorded stimuli presented on a screen can provide particularly powerful settings to study also small effects.

One neural signature that was sensitive to differential social versus non-social attention in typical infants in a naturalistic setting is theta power measured through electroencephalography (EEG). Further, differential theta power responses to social versus non-social stimuli have been shown to be altered in toddlers with autism. Therefore, differential theta power during naturalistic social versus non-social attention might be able to differentiate between groups of infants with diverse outcomes on the level of social interaction skills in childhood. The present study investigated theta power during naturalistic social and non-social attention both in 14-months-old infants with and without elevated likelihood for autism.

3.1.1 Symptoms of interaction difficulties in autism

A condition in which many of the typical signs of social attention are altered is autism, a highly heritable and heterogeneous condition, often co-occuring with other conditions, such as hyperactivity and attention disorders, anxiety and depression (Lord et al., 2022). Worldwide, about 1 in 132 people hold a diagnosis of autism (Baxter et al., 2015). The clinical diagnosis is based on two main criteria (DSM-V, American Psychological Association, 2013), which are deficits in social communication and social interaction across multiple contexts, and at least two circumscribed kinds of current or past restricted, repetitive patterns of behaviour, interests, or activities. Manifestations of autism can further include sensory anomalies and intellectual disability (Lord et al., 2022). Because symptoms might only emerge as social demands exceed the individual's social capacities, a full diagnosis of autism is often only given from three years of age. Early clinical "working" diagnoses can be given to children from two years of age, when behavioural symptoms have started to emerge, in order to track children's further development (Charman & Baird, 2002).

3.1.2 On the origin of autism

Despite its high prevalence, the roots of autism remain unclear. First-degree relatives of individuals diagnosed with autism have a 10-20% higher likelihood of developing symptoms themselves (Constantino, Zhang, Frazier, Abbacchi, & Law, 2010; Ozonoff et al., 2010), suggesting that a genetical component is involved in the origin of autism.

A developmental perspective on autism, supported by an extensive body of literature, suggests that the observable difficulties in social interaction assessed in toddlerhood are manifestations of alterations in neurodevelopmental processes that start much earlier. For typical development, it has been proposed that when children have reached important milestones of social cognition by the end of the first year, social attention has become a function in its own right, independent from non-social attention, and the three attentional subfunctions merge to form the unitary construct of social attention (Braithwaite, Gui, & Jones, 2020; Salley & Colombo, 2016). Within this framework of social attention development, in atypical social developmental, social attention functions might become increasingly dissociated from each other instead of converging to a unifying process. Thereby, some functions may be affected and others not, explaining why in autism some functions of social attention can be relatively intact and others not, the extent to which each function is altered determines the degree to which social attention is present as a unitary construct in the individual. Alterations in general attention or in any component of social attention in specific (social behaviour, social motivation, social attention) may then have cascading effects that may result in social interaction difficulties in toddlerhood (e.g., Dawson, 2008; Johnson et al., 2005; Johnson et al., 2015a). Indeed, this view is supported by evidence indicating that infants who go on to be diagnosed with autism show typical orienting to and engagement with the social world in the first months of life, and then start to gradually withdraw from the social world (Klin et al., 2015; Jones and Klin, 2013; Ozonoff et al., 2010).

On a biological and brain level, symptoms in autism may be the result of developmental interactions between experiences from the environment and neurodevelopmental vulnerabilities (Johnson, 2017). Specifically, the social interaction problems associated with autism are possibly the end result of a cascade of effects caused by adaptive responses to initial widespread mild synaptic dysfunction in early development (Johnson, Jones, & Gliga, 2015). Molecular and genetic factors (but also environmental instability or sensory disturbances) are proposed to affect the balance of glutamergic and GABAergic synapses and cause an excitation/inhibition imbalance. This imbalance leads to poor synaptic fidelity, that is noisier and less reliable transmission of information coming in from the environment. This has particular costs for processing environments that are dynamic and hard to predict such as complex social interactions.

Johnson and colleagues (2015) argue that while this early atypicality in synaptic functioning may be mild and transient, it is diffuse and widespread and can further cause a series of compensatory and adaptive processes, with cascading effects on ongoing development. For example, since the dysregulation might affect some regions of the brain more than others, individuals may compensate performing functions of affected brain systems by recruiting alternative brain systems. Also, an excitation/inhibition imbalance might account for increased local connectivity (to reduce noise by averaging information) and decreased long-distance connectivity (to prevent additional noise by decreasing the number of synaptic connections the information needs to cross). Further, cortical specialisation might need longer as a result of the unreliability of the processed input from the environment. Also, individuals might adapt their attention style to fit their processing style. For example, poor reliability in sensory processing may be multiplied in multi-modal sensory integration and limiting sensory input by employing overly focussed attention can help reduce the noise. The unreliability in incoming information may reduce the reward infants experience when processing social information, subsequently adapting to orienting more towards predictable, repetitive events such as mechanical, self-controlled stimulation for which processing has proved more successful and rewarding than for social interaction. Also, with development, more possibilities arise for actively seeking information in the more predictable non-social environment, such as motor skills or the ability of attention control. Reduced orienting to and engagement with the social world in turn reduces the amount of social experiences, resulting in atypical cortical specialisation of the social brain network. These kinds of typical adaptations to dysfunctional synaptic functioning then manifest in atypical behaviours associated with autism.

Under this framework, to understand the pathway of emergence of social behaviour difficulties characteristic to autism requires studying the trajectories of social development from infancy through toddlerhood, and to identify early signs of divergence (Jones, Gliga, Bedford, Charman, & Johnson, 2014). Identifying early markers on the level of looking behaviour or neural processes, which mediate between genetic influences and behavioural symptoms, would help understand protective and risk factors in the path to behavioural problems and may facilitate an earlier diagnosis (Braithwaite, Gui & Jones, 2020). Specifically, studying the origin of autism in its roots requires studying when synaptic dysfunction occurs, which types of information processing are affected by it and which parts of the brain (Johnson et al., 2015). Besides understanding which mechanisms are involved in the developmental cascades leading to diagnosis, the identification of early biomarkers would pave the way for interventions to take effect at a stage where substantial change is still happening and where interventions may be able to break the developmental cascade towards social behaviour difficulties. Reliable markers are also needed to evaluate the success of interventions by allowing to reveal potentially small effects. Detecting also small effects is important, since they might result in large cascading effects down the line (Jones, Dawson, Kelly, Estes & Webb, 2017). Of note, Johnson et al. (2017) point out that the altered attention style has an adaptive function that serves the individual to navigate easier in their environment and increase their quality of life. Early interventions targeting the initial synaptic dysfunction would be most beneficial in preventing the social interaction difficulties associated with autism. Increasing predictability in the environment, either in the (social) environment itself or by orienting to more predictable aspects in the environment, might help individuals to reduce effects of inconsistent processing.

Symptoms of ADHD, with attention problems not specific to the social domain and often emerging later than autism symptoms, might be the end result of adaptive responses to similar neural dysfunction but expressed at a later time point in development, when social and non-social attention have diverged into separate constructs (Johnson, Gliga, Jones, & Charman, 2015).

With regards to theta power, disrupted synaptic functioning and subsequently lower reliability of incoming information, coordinated theta activity might be decreased as a result of disrupted synaptic transmission. As discussed in Chapter 1, long-term potentiation (LTPs) represents synaptic plasticity that facilitates signal transmission across neurons and has shown to facilitate memory and learning. LTPs more likely arose when incoming stimulus sequences in the hippocampus happened at the theta frequency compared to other frequencies (Greenstein et al., 1988). Reduced fidelity in coordinated synaptic activity and hence information transmission across neurons, which may be particularly impaired in the social domain where a multitude of unpredictable multimodal cues have to be integrated, might result in a reduction of the typical enhancement in theta rhythms arriving at the scalp that has been observed during social versus non-social information processing in autism (Dawson, Bernier, & Ring, 2012).

To sum up, social interaction difficulties may be the consequence of typical, adaptive, helpful responses of the brain to early widespread synaptic dysfunction, mediated by markers on the level of brain processing such as a reduction in the typically observed elevated theta power during social compared to non-social processing.

3.1.3 Studying the emergence of social behaviour symptoms: The prospective design

Studying the early onset of autism requires investigating individuals in infancy when processing differences are first emerging. To do this, research has made use of the high heritability that characterises autism. Studying children or siblings of diagnosed individuals in infancy implies a 10-20% chance of studying an infant with later autism diagnosis. Responses are compared with responses of infants with typical likelihood of autism who do not have a first-degree relative holding an autism diagnosis. The prospective design follows up the development of these infants with elevated likelihood over multiple timepoints up into childhood age when they are clinically assessed, and a diagnosis can be made. Therefore, the prospective design allows to study early differences in infants with elevated likelihood for autism, including those who went on to be diagnosed.

3.1.4 Signs of atypical social attention in early autism

Studies using the prospective design have produced evidence supporting the above framework of atypical social brain development (for a review see Jones et al., 2014).

3.1.4.1 Behavioural signs of atypical social attention in early autism

On the behavioural level, infants with elevated likelihood for autism showed shorter peak look durations overall and later peak looks to faces but not objects compared to typical infants at 6 and 12 months (Jones et al., 2016). At 6 months of age, infants with later diagnosis of autism spent less time looking at the human in a video than infants without later diagnosis, and of the time looking at the human, they spent a smaller proportion of time looking at the face than typical infants (Chawarska, Macari, & Shic, 2013). At 7 months, infants with later autism diagnosis did not differ in visual disengagement from central stimuli, but they took longer to disengage from the stimulus by 14 months (Elsabbagh, Fernandes, et al., 2013). Moreover, between 7 months and 14 months, infants with later diagnosis did not show the typical increase in visual orienting speed and flexibility (Elsabbagh, Fernandes, et al., 2013).

Further, at 10 months, infants with elevated likelihood showed less frequent gaze-based initiation of *joint attention* with a live actor but typical amounts of overall looking (Nyström, Thorup, Bölte & Falck-Ytter, 2019). Similarly, at 10 months, infants with elevated likelihood for autism showed different rates of responding to joint attention measured by gaze following (Nyström et al., 2019) and less looking at the adult in the second after the adult initiated direct gaze (but not beyond that second) (Nyström et al., 2017). At 7 and 13 months, infants with elevated likelihood did not differ in gaze following (Bedford et al., 2012).

At 12 months, infants with later autism diagnosis did not look more at the face during parent-child interaction in a live interaction setting, while this was the case for typical infants (Gangi et al., 2018). Similarly, 6-month-old infants with later autism diagnosis looked less at the experimenter's face than infants without later diagnosis during structured cognitive assessment by an experimenter (Mullen Scales of Early Learning Mullen, 1995; Gangi et al., 2021). Further, a systematic review of 15 studies on parent-infant interaction in infants with elevated likelihood for autism using the prospective design indicated differences in parent-child interaction in infants with elevated likelihood for autism, primarily on the level of gestures and dyadic interaction quality, albeit there was no relation with later diagnosis (Wan, Green & Scott, 2019; but note Pijl et al., 2021 for no differences).

As already mentioned, infants with later autism diagnosis displayed an atypical *decline* in the trajectory of overt attention to social stimuli across development. While this has been observed consistently across studies, different timepoints of the onset of this decline have been reported. Attention to the eyes has been reported to decline in infants with later autism diagnosis between 2 and 24 months (Jones & Klin, 2013) and between 6 and 12 months after appearing typical at 6 months (Ozonoff et al., 2010). From 12 through 36 months of age, infants with later autism diagnosis showed a sharper decline than typical infants in the proportion of looking at the face during cognitive assessment with a researcher and during parent-child-interaction (Gangi et al., 2021). This pattern is in line with the view that social attention may be normal at first but then decline reflecting adaptations to disrupted synaptic functioning.

Besides social processing, behavioural differences have also been found for *non-social* information, from birth up to toddlerhood. A recent study reported differences in non-social processing as early as a few days after birth. Specifically, newborns with elevated likelihood for autism showed more looking at inverted face-like patterns, a higher number of looks and longer look durations at a random motion stimulus compared to typical newborns (Di Giorgio et al., 2016). Altered responses to overall social and non-social were also observed at 3 months, as well as a relation between *overall* attention at 3 months and developmental outcome at 12 months observed across both groups (Bradshaw et al., 2020). Finally, non-social processing was altered in 14- to 42-month-old toddlers with autism. They spent a larger proportion of looking at a video of geometric patterns versus a video of children in action, and the proportion of looking at geometric patterns predicted autism diagnosis status (autism/no autism) (Pierce, Conant, Hazin, Stoner & Desmond, 2011).

3.1.4.2 Neural signs of atypical social attention in early autism

Alterations in differential processing of social versus non-social information in infants with later autism diagnosis were further shown on the neural level. For example, infants with later autism diagnosis showed a smaller and shorter Nc response to faces at 6 months (Jones et al., 2016). Similarly, at 8 months, infants with elevated likelihood for autism showed a smaller Nc response to faces with direct gaze and a larger Nc response to noise, while larger Nc amplitudes to noise were related to poorer social behaviour outcomes at 3 years (Gui et al., 2021). Infants with later autism diagnosis further showed alterations in other ERPs involved in face processing in response to gaze shift at 6-10 months (Elsabbagh et al., 2012). Further, 8-month-old infant siblings showed altered ERP responses to gaze shifting towards versus away from the viewer as well as to face versus noise stimuli (Tye et al., 2020), predicting autism diagnosis among infant siblings at elevated likelihood at an individual level with ~77% accuracy (Tye et al., 2020). However, there is also evidence reporting no difference in ERP responses between infants at elevated and typical likelihood for autism when averaged across the time between 6 and 36 months as well as at 12 months (Luyster, Powell, Tager-Flusberg & Nelson, 2014).

Processing differences relating to autism have also been observed on the level of theta power in more naturalistic contexts, with autistic toddlers showing reduced elevation of theta power to social versus non-social videos (Dawson, Bernier, et al., 2012). However, in a recent study with 14-month-old infants watching social versus non-social videos, theta power did not differ significantly between infants with typical and elevated likelihood for autism; the typical effect of stronger theta power during social versus non-social processing was observed across groups (Haartsen et al., 2022). Also, theta power differences were not significantly related to dimensional measures of social behaviour at 3 years of age (Haartsen et al., 2022).

Differences have also been observed using other techniques. For example, extracting typical microstates reflecting attention to faces allowed to reveal a relation between microstate duration and autism diagnostic outcome, and between microstate strength and social skills outcomes (Gui et al., 2021). Further, the inter-trial consistency in the theta band recorded during face processing was reduced in infants with later autism diagnosis (Van Noordt et al., 2022). On the level of activation of brain regions, infants around 5 months who went on to be diagnosed with autism showed reduced social brain activation to auditory (Blasi et al., 2015; Lloyd-Fox et al., 2018) and visual (Lloyd-Fox et al., 2018) social versus non-social stimuli. Further, at around 5 months, infants at elevated likelihood for autism showed reduced activation to visual (Braukmann et al., 2018; Lloyd-Fox et al., 2013) and auditory (Lloyd-Fox et al., 2013) social stimuli.

Together, these findings suggest that there are early behavioural and neural differences in infants who go on to be diagnosed with autism that are cascading from reduced looking at the eyes in early infancy to reduced attention to the social partner over the course of the first year, while there is also evidence of typical responses towards social stimuli in infants with later autism diagnosis, as well as altered responses towards non-social stimuli. This pattern supports the above proposal that social and nonsocial may be the same first and diverge over early development, that different functions of social attention may be differently affected in atypical development, and that brain specialisation to social and non-social stimuli might both be altered as a consequence of initial synaptic dysfunction and its cascading effects on ongoing brain development, with the extent of both alterations depending on the timepoint of the dysfunction.

Besides findings suggesting early differences in visual social attention, primarily studied in relation to face processing, there is also evidence of differences in auditory processing. A meta-analysis revealed that young children with autism tended to show a smaller mismatch negativity amplitude than typically developing children (Schwartz, Shinn-Cunningham & Tager-Flusberg, 2018). Further, 14-month-old infants at elevated likelihood of autism showed a reduced ERP response towards their own name versus an unfamiliar name compared to typical infants (Arslan et al., 2020).

While numerous studies have pointed out early differences in behavioural and neural processing in infants with elevated likelihood for autism and/or who go on to be diagnosed themselves, no study to date has compared these groups in a naturalistic social context. This suggests that social processing differences in early autism go beyond the boundaries of single modalities and hence require to be studied in multimodal contexts.

3.1.5 Differential theta power as sign of atypical attention during naturalistic social experiences

One neural signature that has been revealed to be sensitive to multimodal social contexts is theta power, which was stronger during social versus non-social live action in 6- and 12-month-old infants (Jones et al., 2015). Beyond fostering our understanding of typical development, theta power differences have been revealed in adults and young children diagnosed with autism. For example, adults with compared to without Asperger syndrome (n=5) showed lower levels of whole-scalp theta power increase in time-logged responses to faces (n=7) (Yang, Savostyanov, Tsai & Liou, 2011). Similarly, theta power over central and posterior regions increased with longer looking duration at social compared to non-social stimuli only in typically developing children, while it decreased in children with autism (Isaev et al., 2020). Further, theta power (5-7 Hz) towards faces versus objects was reduced in toddlers (aged 48-77 months) with autism compared to typically developing toddlers (Dawson, Bernier, et al., 2012). Finally, children with autism who had received a 2-year behavioural intervention previously shown to improve social behaviour skills exhibited typical patterns of theta power modulation for faces versus objects (Dawson et al., 2012), highlighting the tight coupling between theta power and social information processing. Promisingly, behavioural intervention was able to attenuate theta power alterations in infants with elevated autism likelihood: Nine-to-eleven-months-old infants at elevated likelihood for autism received parent-delivered intervention that had previously been shown to improve parental responsivity to infant social communicative cues. These infants exhibited greater increase in frontal theta power from 6 to 18 months in response to social versus non-social videos as well as a more typical response pattern during watching social versus non-social videos compared to infants receiving a control intervention (Jones et al., 2017).

Together, these findings indicate that stimulus-driven theta power modulation is altered in individuals with autism, from toddlerhood through adulthood. Because theta power, especially over frontal cortical regions, has been shown to reflect top-down control processes such as sustained attention during infancy, and is modulated by the social content of a situation in typically developing infants, it may be sensitive to altered social processing in early autism before behavioural symptoms emerge. However, no known published study to date has investigated whether social versus non-social theta power during naturalistic social experiences is different in infants with elevated autism likelihood and with later autism diagnosis.

3.1.6 Specificity to autism: Theta power in ADHD

Like autism, ADHD is a common neurodevelopmental condition. ADHD is usually diagnosed in childhood around early school age (Sainsbury et al., 2023) and it is characterised by inattentiveness and/or hyperactivity and impulsiveness (DSM-V; American Psychiatric Association, 2013). Autism and ADHD often cooccur, with 30-80% of the children diagnosed with autism also meeting criteria for ADHD (Rommelse, Franke, Geurts, Hartman, & Buitelaar, 2010). However, it is not clear whether their causal pathways are the same or distinct, that is whether they are different manifestations of the same underlying condition, or whether ADHD symptoms are a secondary consequence of autism symptoms (Shephard, Bedford, et al., 2019).

There is scarce research about theta power in relation to ADHD traits and diagnosis in children. During resting state, frontal theta power has been repeatedly shown to be stronger overall in children with compared to without ADHD (Barry, Clarke & Johnstone, 2003; Tye, Rijsdijk, & McLoughlin, 2014). Further, in children with ADHD, theta power was enhanced over frontal electrodes in response to a cue (still image of a pair of human eyes) that indicated the soon onset of a target stimulus (Guo et al., 2020). However, theta power recorded from 10-month-olds while watching social and non-social videos was not related to ADHD traits (Goodwin et al., 2021), suggesting that the overall elevated theta power might not be modulated by the social nature of a stimulus.

3.1.7 Summary

Autism is a neurodevelopmental condition characterised by social behaviour difficulties from toddlerhood on. The developmental roots of autism are still unclear. Evidence supports views proposing an early alteration of the ability to process basic social information, resulting in a cascading development towards the visible behavioural manifestations at toddler age, when social demands exceed capacities. One of the questions that remain unsolved in the field is the question about robust early markers for autism, which could be used on the individual level to predict outcome and design interventions. While a recent study showed no difference in theta power responses during social versus non-social video viewing in 14-month-old infants with elevated likelihood for autism, live contexts have proven more powerful than screen-mediated stimuli in triggering differential theta responses in typical infants, and might by their complex dynamic nature potentiate processing difficulties of social cues in early autism. Thus, they might be particularly suited to reveal potential differences between groups. Theta power has been shown to be a possible candidate for revealing early processing differences in infants with later autism diagnosis, due to its increasing sensitivity towards social stimuli over the first year of age and its differential response to social stimuli being altered in toddlers with autism.

3.1.8 The present study

While a substantial amount of literature reports stronger theta power during social compared to non-social attention in the first year of age (e.g. Jones et al., 2015), response patterns seem to be altered in toddlers with autism (Dawson, Bernier, et al., 2012) and might already be in infancy. To date, however, no studies have measured theta power during naturalistic social and non-social stimulation in elevated-likelihood infants. Towards analysing individual brain responses in real-time during naturalistic social experiences, this chapter aims at investigating whether theta power during social versus non-social naturalistic experiences differs in infants with elevated likelihood of autism.

In the present study, 14-month-old infants with and without familial history of autism and/or ADHD were viewing a live experimenter singing nursery rhymes (social condition) and operating a dynamic toy (non-social condition), respectively, while EEG was recorded. The study aimed to investigate whether and how previously observed theta power modulation by live social versus non-social stimulation is different in infants with elevated likelihood for autism. The age of 14 months was chosen because this is when behavioural symptoms in infants with atypical social development first start to emerge (Jones, Gliga, Bedford, Charman & Johnson, 2014). Theta power activity was measured in relation to stimulus condition and group. Theta power has been suggested to reflect attention engagement and learning, and is elevated during social versus non-social processing in typical infants (Jones, Venema, Lowy, Earl & Webb, 2015), while theta modulation by social context was reduced in toddlers with autism (Dawson, Bernier, et al., 2012). Live social stimuli were used to increase ecological validity and because theta modulation by social stimuli was more pronounced in a live social context (Jones et al., 2015).

Theta power was predicted to be stronger in the social versus non-social condition in the overall sample. Second, this effect of differential theta power was predicted to be reduced in infants with elevated likelihood for autism. Third, the proportion of looking at the stimulus was not predicted to be different in infants with elevated likelihood for autism. Fourth, the effect of differential theta power was predicted not to be reduced in infants with elevated likelihood for ADHD. Fifth, the effect of differential theta power was expected to be reduced in relation to autism diagnosis and higher scores on autism trait measures at 36 months. Sixth, the effect of differential theta power was expected not to be related to ADHD trait measures at 36 months. Finally, the effect of differential theta power at 14 months was expected to predict the trajectory of looking at a face at 10, 14 and 24 months.

3.2 Methods

3.2.1 Participants

Infants were part of the British Autism Study of Infant Siblings (BASIS, www.basisnetwork.org), a longitudinal study investigating development of infant siblings of children diagnosed with autism and/or ADHD. The cohort that provided data for the present study consists of 166 infants (Phase-3 cohort, *STAARS, Studying Autism and ADHD in the eaRly yearS*, 2013-2019). Included in the study were infants born full-term (gestational age greater than 36 weeks) and who did not have a known medical or developmental condition at the time of enrolment. Participants were recruited over the BASIS study website and the recruitment portal of the Birkbeck Babylab (*Centre for Brain and Cognitive Development, University of London*). Families who had enrolled in the study were invited to visits when their baby was 5, 10, 14, 24 and 36 months old.

The current analysis focused on EEG data obtained at 14 months. In total, 100 infants provided data from the EEG live-singing task at that timepoint. The remaining 66 infants of the cohort had to be excluded from the present analysis because they missed this timepoint visit (n = 24), the child refused to wear the EEG net or was too upset (n = 29), because the paradigm was not administered because no time was left during the visit (n = 4), because the child fell asleep (n = 3), because the child had a chin rash (n = 1), due to technical failure (n = 2), because no video was recorded during the EEG task (n=1) or because the EEG data file went missing (n = 2).

3.2.2 Autism and ADHD likelihood assessment

Infants entering the BASIS study either had a first degree relative (i.e. parent or sibling) with autism diagnosis, with ADHD diagnosis or probable ADHD, or neither of the two. The presence of autism in the first degree relative was attested through a clinical autism diagnosis by a licensed clinician. Elevated likelihood for ADHD (i.e. the presence of ADHD in the first degree relative) was assigned either if a first degree relative had a clinical diagnosis of ADHD, or if despite the absence of a clinical diagnosis concerns of ADHD traits in the family had been reported and subsequently confirmed in screening assessment administered by BASIS researchers at the lab. A short version of the Conners Early Childhood 3 form (Conners, 2008) was used as a screening assessment for siblings younger than six years. A shortened version of the Conners 3 was used for screening assessment of first-degree relatives aged six or older. This second route of assigning ADHD likelihood status was included, because ADHD is often not diagnosed in children who already hold a diagnosis of autism (e.g., Visser, Rommelse, Greven & Buitelaar, 2016). Therefore, this additional route was applied to prevent under-identification of ADHD conditions in autism-diagnosed children. Infants with typical likelihood of autism or ADHD were recruited through the volunteer family database hosted by the Centre for Brain and Cognitive Development, Birkbeck University of London. These infants had no first-degree relatives with a diagnosis of autism or ADHD, and at least one older sibling with typical development. Each infant was assigned a value on a dichotomous variable for typical (0) or elevated (1) likelihood for autism, and on another dichotomous variable for typical (0) or elevated (1) likelihood for ADHD.

The sample of 100 infants with valid EEG data included infants with typical autism likelihood (n = 14), elevated autism likelihood (n = 53), elevated ADHD likelihood (n = 18) or elevated likelihood for both conditions (n = 13). Of note, analyses were run with factors of autism presence or absence as well as ADHD presence or absence.

3.2.3 Procedure and measures

Participants of the BASIS study visited the Babylab several times within their first years of age. Each visit consisted of a battery of eye tracking, behavioural and neurophysiological tasks. There caregiver gave informed written consent before the start of the study. The testing was only conducted if the infant was being content and alert. After each visit, families were given a participation certificate and t-shirt. If required, expenses for travel, subsistence and overnight stay were reimbursed. The protocol of this study was approved by the *National Research Ethics Service* and the *Research Ethics Committee of the Department of Psychological Sciences, Birkbeck, University of London*. Each of the assessments and experiments during the visits were performed in accordance with the relevant guidelines and regulations. Data were collected by the STAARS team.

3.2.4 EEG task

The same EEG task was administered at the 5-, 10- and 14-month visit. The current analysis focusses on the 14-month-timepoint. In the EEG task, each infant was confronted with a new experimenter. This experimenter displayed either socially engaging or socially non-engaging behaviour in blocks of one minute, respectively. The blocks were alternating until three blocks for each condition were reached. In the social blocks, infants watched the experimenter singing nursery rhymes accompanied by gestures, direct gaze and friendly, engaging facial expressions. In the non-social blocks, infants watched a dynamic toy spinning on a low table just out of reach for the infant, operated by the same experimenter. For consistency with previous studies into naturalistic social versus non-social processing, the terms "social" and "non-social" condition to refer to the condition including the singing and the spinning toy, respectively. Interaction between experimenter and infant was reduced to a minimum. The parent was asked to not engage with the infant, but to react as usual if the infant would initiate the interaction. Infants sat on their parents' lap in an electrically shielded room. The experimenter sat on a chair opposite the child. EEG was recorded at a sampling rate of 500 Hz using a 128-channel HydroCel GSN 130 EGI electrode net and Netstation EGI software. Infant behaviour was video recorded for later coding. The session was stopped when the infant became uncomfortable or fussy.

3.2.5 Eye tracking task (10-, 14- and 24-month visit)

In the eye tracking task ("Fifty faces" task, Vo, Smith, Mital, & Henderson, 2012), a video clip (1280 px X 720 px, 25 fps, 41 seconds) from the Dynamic Images and Eye Movements database ("Fifty People One Question: Brooklyn", http://thediemproject.wordpress. com/) was presented to the child, in which various pedestrians in New York were recorded on camera while speaking to the camera for an interview. The clip represented a natural scene of social interaction with actors naturally starting and stopping talking and making and breaking eye contact. Subtle music was used to replace the speech track in order to avoid confounds with language. Stimuli were presented using custom Matlab scripts (Task engine, https://sites.google.com/site/taskenginedoc), and infants' eye movements were recorded with a Tobii TX-300 eye tracker. Areas of interest (AOI) were drawn on each single video frame using Apple Motion, including the face AOI as well as a body AOI and a background AOI. The face AOI followed the outline of the face, below the hairline and along the jawline. Infants' gaze samples were assigned to AOIs for each frame. If there was a gap <200ms between the same AOI, this gap was interpolated by the same AOI. Data was excluded if there were fewer than 25% valid samples (i.e. looking to no AOI) for the video. The proportion of looking time to all AOIs was calculated as the number of samples per AOI by the total number of valid samples for the video. Four measures of visual attention to faces were coded and used in the present analysis: (1) the *proportion of time* the infant spent looking to the face AOI, (2) the total *number of looks* to the face AOI, (3) the *mean fixation duration* to the face AOI, and (4) the *peak fixation duration* to the face AOI (i.e., the longest sequence of samples at one AOI).

3.2.6 Autism outcome assessment

Autism outcome was assessed at the 36-month-old visit of the BASIS study. Phenotypical characteristics of autism were measured using the Autism Diagnostic Observation Schedule Toddler Module (ADOS-T; Luyster et al., 2009) and the ADI-R, and parents filled in the Vineland Adaptive Behaviour Scale II (VABS-II; to obtain the Socialisation Standard Score) as well as the Social Responsiveness Scale (SRS-2 Preschool; to obtain the Total T-score). Further, at 36 months, categorical outcome was clinically assessed (autism yes/no).

3.2.6.1 Autism Diagnostic Observation Schedule – T

At the 36-month-timepoint, the Autism Diagnostic Observation Schedule Toddler Module (ADOS-T; Luyster et al., 2009) was administered. The ADOS is a semi-structured, standardised observational play based behavioural assessment that measures autism traits. It measures joint attention, eye contact, focused attention, expression of emotions, language use, and others. In the Toddler Module, the items are presented in a playful manner by an experienced clinician. The frequency, quality, and appropriateness of toddlers' responses to the items as observed during the session are rated. Item scores are combined into domain algorithm scores. The current study includes the Social Affect domain score measuring communication and reciprocal social interaction. I used Calibrated Severity Scores because these consider age and language level of the child while they are based on the raw total scores of the ADOS (Esler et al., 2015). Higher scores on the scales indicate more severe symptoms.

3.2.6.2 Autism Diagnostic Interview – Revised (ADI-R)

At the 36-month-timepoint, the Autism Diagnostic Interview – Revised (ADI-R, Rutter et al., 2003) was additionally used to assess symptom severity. The ADI-R is a standardized, semi-structured interview that is conducted with the parent or caregiver of the child. It consists of 93 items that investigate both past and current severity of autism symptoms, the onset of the symptoms, the acquisition and loss of language and other skills, language and communication functioning, social development and play, interests and behaviours, and other general behaviours such as gait, aggression or special skills. Item scores are combined into domain algorithm scores. The present analysis includes the ADI-R Social Total score. Higher scores reflect more severe symptoms.

3.2.6.3 Parent report measures

At the 36-month-timepoint, parents were asked to provide information on their child's behaviour.

The **Vineland Adaptive Behaviour Scales II** (VABS-II; Sparrow, Cicchetti & Balla, 2005) is a semi-structured interview measuring adaptive functioning in everyday life (Sparrow et al., 2005). It has been used with infants and toddlers to capture differences in developmental trajectories (e.g., Bussu et al., 2018). The present thesis uses the questionnaire form of the scale (Parent/Caregiver Form). In the present analysis, the Socialisation domain of the VABS-II was included to relate neural differences in social versus non-social processing to adaptive behaviours in daily social situations. Higher scores represent more socially adaptive behaviours. The overall Socialisation standard score was used as well as the two subdomain scores of "Interpersonal Relationships" and "Play and Leisure Time". Vineland v-scale scores have a mean of 15 and a standard deviation of 3. Scores range from 1 to 24. Higher scores represent more adaptive functioning.

The **Social Responsiveness Scale** (SRS, Constantino & Gruber, 2012) is a continuous measure of social ability. It can be used to assess autism phenotypic trait variation by capturing variation of social impairment in undiagnosed individuals. Higher scores represent more severe impairments. Of note, higher impairment measured by the SRS (SRS-2 Preschool Total at 36 m) was previously found to be associated with reduced hemodynamic responses to visual social stimuli in 4- to 6-month-old infants with later autism diagnosis (Lloyd-Fox et al., 2018).

In addition, at the 36-month-timepoint, ADHD outcome was assessed using the **Child Behaviour Checklist** (CBCL, Achenbach and Rescorla, 2001). The present analysis includes the ADHD scale (total T score) of the CBCL. This scale measures inattention and hyperactivity in toddlers by asking parents to rate how well their child's behaviour matches with each statement. Parents are asked to indicate on six items how well each statement describes their child's behaviour as observed within the past 2 months on a 3-point Likert rating scale, from 0 ("Not True") to 2 ("Very True or Often True"). Item scores are then summed to produce a total score. In the present analysis, missing items of the CBCL are treated as 0 ("Not True"), as recommended in the manual of the checklist. Higher scores reflect more severe symptoms.

3.2.6.4 Categorical outcome (autism yes/no)

At the 36-month-timepoint, categorical outcome (Autism / No Autism) was assessed by a team of clinical researchers lead by a licensed clinical psychologist after reviewing all relevant information available up to and including 36 months. Relevant information included the results from the ADI-R, ADOS-2, and the VABS-II.

3.2.7 EEG pre-processing and analysis

3.2.7.1 EEG pre-processing

EEG data storage and pre-processing were carried out by Jannath Begum Ali in EGI NetStation 4. The video recordings were coded second-by-second to determine the segments during which the child was looking at the stimulus. Only trials in which infants were looking at the respective stimulus (experimenter in the social condition, toy in the non-social condition) entered the analysis. In the social condition, looking was coded as *looking at the experimenter / not looking at the experimenter*. In the toy (= non-social) condition, looking was coded as *looking at the toy / not looking at the toy*. Segments in which the infant was crying or hiding their face or in which the experimenter left the video were categorised as not valid.

The EEG recording was digitised with a sampling rate of 500 Hz, filtered with 0.1 high-pass and 100 Hz low pass filter, and segmented into 1 second segments. Quality of each channel was assessed by segments of one second. Within each segment, bad channels were identified and excluded. Channels were classified as bad if they showed an exceptionally high amplitude or frequency, if they differed a lot from the signal in the neighbouring channels or if their signal was drifting in more than one direction. The signal of excluded channels was interpolated by the signal of the neighbouring channels, by an algorithm incorporated within NetStation 4.3. The whole 1-second segment was excluded if it contained more than 25 interpolated channels in total or contained clusters of interpolated channels (i.e. >6 channels next to each other). Infants were excluded due to bad data quality if there remained less than 10 segments in any of the two conditions after the cleaning process. Data were re-referenced to the average. The resulting segmented data was imported into MATLAB.

3.2.7.2 EEG spectral power analysis

EEG spectral power analysis was carried out by Emily J. H. Jones. In-house MATLAB scripts were used to process a Fast Fourier transformation on the data revealing power spectra. A channel was excluded from a given segment if power values exceeded three standard deviations from the mean of that frequency in the remaining channels in that cortical region. Power values were averaged across artifact-free segments and across electrodes within each cortical region (frontal, parietal, occipital, temporal).

The term “power” here refers to the absolute power spectral density (aPSD), that is the power present in the EEG signal as a function of frequency, per unit frequency.

“Theta power” accordingly is the power in the infant EEG signal in the theta frequency band (3-6 Hz).

To obtain relative power values, natural absolute power values were averaged across the theta frequency range and divided by the total signal power for each region over each brain side in each condition. In specific, for each of four cortical regions (frontal, parietal, occipital, temporal) and each of three possible sides in the brain (left, central, right) in each condition (social, non-social), relative theta power was calculated as the sum of the exponential function of the absolute power at each single Hertz unit from 3 Hz to 5.9 Hz, divided by the total signal power, that is the sum of the exponential function of the absolute power at each single Hertz unit from 3 Hz to 20 Hz. This is expressed in the formula for relative theta power:

$$\mathrm{relTheta}_{\mathrm{region1,\ condition1,\ side1}} = \frac{\mathrm{sum(exp(aPSD3Hz)+exp(aPSD4Hz)+exp\ (aPSD5Hz))}}{\mathrm{sum(exp(aPSD3Hz)+exp(aPSD4Hz)+\ \ldots\ +exp\ (aPSD20Hz)}}$$

3.2.8 Statistical analysis

The statistical analysis was carried out in R (R 3.6.2, R Core Team, 2019). The following models were calculated:

1. Likelihood

   A mixed model ANCOVA was used with relative theta power as dependent variable and stimulus condition (social/non-social) as within-subject factor and autism likelihood (TL-autism/EL-autism), ADHD likelihood (TL-ADHD/EL-ADHD) and their interaction as between-subject factors. The model was re-run with the proportion of looking to the stimulus as dependent variable.

2. Autism diagnosis outcome (yes/no) at 36 months

   A mixed model ANCOVA was run with relative theta power as dependent variable, including stimulus condition (social/non-social) as within-subject factor and autism outcome (no autism/autism) as between-subject factor. The model was re-run with the proportion of looking to the stimulus as dependent variable.

3. Autism-related traits at 36 months

   A linear mixed model analysis was run with relative theta power as dependent variable. Fixed effects were stimulus condition and four continuous measures of autism symptoms (ADOS; ADI social scale; SRS total T score; Vineland Socialisation standard score) as well as their interactions with stimulus condition. Age in days (centred), gender and also the number of valid EEG segments were added as covariates, and participant was included as random effect. The model was re-run with the proportion of looking to the stimulus as dependent variable (and with only age in days and gender as covariates).

4. ADHD-related traits at 36 months

   A linear mixed model analysis was run with relative theta power as dependent variable. Fixed effects were stimulus condition and the dimensional ADHD traits measured at 36 months on the CBCL as well as the interaction with stimulus condition. Age in days (centred) and gender, as well as the number of valid trials were added as covariates, and participant was included as random effect. The model was re-run with the proportion of looking to the stimulus as dependent variable (and with only age in days and gender as covariates).

5. Trajectory of looking at the face

   I ran four linear mixed models were run to investigate whether differential theta power during social processing in infancy predicts the trajectory of looking at a face across toddlerhood. In each model, the dependent variable was one of the following characteristics of face looking: a) the proportion of time, b) the number of looks, c) the peak look duration, and d) the mean look duration. Fixed effects in each model were relative theta power at 14 months, the proportion of looking to the live stimulus at 14 months, as well as stimulus condition (social/non-social), the timepoint of when the outcome eye tracking measure was obtained (10/14/24 months) and their interactions with the theta power and the proportion of looking. Age in days (centred), gender, and

> number of valid EEG segments were controlled by including them as additional fixed effects in the models. Participant was included as random effect.

Where significant effects on theta power were observed, the data were further analysed in respect to electrode side (left, central, right) and cortical region (frontal, parietal, posterior, temporal). The linear mixed models were computed with the *lme4* R-package (Bates, Sarkar, Bates, & Matrix, 2007). The number of valid EEG trials was controlled for to account for potential differences in the amount of looking to the social versus non-social stimulus and the resulting unequal number of valid EEG trials available for the two conditions. To this end, the standardised difference of the number of EEG trials was calculated for each participant ($Z = (x-\mu)/\sigma$) and included this score as a covariate in the models. Missing covariate values (age and trial number) were replaced by the mean of all covariate values across the sample.

Testing model assumptions for the mixed model ANCOVAs included the assessment of outliers in each cell of the design, based on visual inspection of boxplots and using the identify_outliers function of the R rstatix package (0.6.0, Kassambara, 2020). Extreme outliers were reported. Further, for each cell of the design, normality was assessed visually by inspecting QQ-plots. If the dependent variable was non-normal, a transformation was applied so that it after normalization followed a Gaussian distribution. The specific transformations applied, respectively, were selected by the bestNormalise package in R (Peterson & Cavanaugh, 2019), that helps identify the best normalizing transformation for a given vector. Transformations used were the Ordered Quantile (ORO) normalization transformation ("orderNorm") transformation and the Yeo-Johnson transformation ("yeo_johnson"). The former normalizes a vector by mapping the values to their percentile and maps this to the same percentile of the normal distribution. The latter is a more complex transformation allowing the vector to also include zeroes and negative values. Finally, homogeneity of variances was assessed. To do this, I combined Levene test results with visual inspection (hovPlot function, *HH package*, Heiberger, 2020), because the number of data points is relatively high and thus prone to positive results of the Levene tests even if the variances are relatively homogeneous. Additional ANCOVA-specific assumptions included that on each level of the categorical variables the covariate was linearly related to the dependent variable and the slope of the relationship between the covariate and the outcome was the same across conditions. Testing model assumptions for linear mixed models included checking that the group variances are homogeneous, that the residuals of the model are normally distributed and that residuals are independent from the predicted values.

For each analysis, main effects and interaction effects that include stimulus condition are reported first since these were or primary interest for my hypothesis. To follow-up significant interactions, I used separate ANCOVAs or linear mixed models for each stimulus condition. Main effects and interactions not including stimulus condition are reported but were not investigated in more depth. Infant numbers per likelihood group and by outcome measures are listed in Table 3.1 and Table 3.2, respectively.

Of note, autism and ADHD likelihood were included as separate factors with either typical or elevated likelihood, respectively. This design allows to test main and interaction effects of both factors, while it does not allow to directly compare effect sizes.

Table 3.1. Number of infants with valid EEG data at 14 months, by likelihood group; *these additional infants with valid EEG data at 14 months were excluded because they were half-siblings or had unclear likelihood status

Likelihood group	Total	TL	EL-ASD	EL-ADHD	EL-both
Total cohort	N=161(+5*)	29	80	31	21
Available EEG data at 14 m	n = 98(+2*)	14	53	18	13

Table 3.2. Number of infants providing data on the tasks and measures used in the present study. ADOS-2 = Autism Diagnostic Observation Schedule; ADI = Autism Diagnostic Interview; VABS-2 = Vineland Adaptive Behaviour Scales; SRS= Social Responsiveness Scale; CBCL = Child Behaviour Checklist

	Timepoint				
	5 months	10 months	14 months	24 months	36 months
EEG Live task	(n = 92)		n = 100	-	-
Eye tracking “50 Faces” task	-	n = 115	n = 124	n = 91	-
Diagnostic assessment	-	-	-	-	n = 128
ADOS-2 Social Affect (CSS)	-	-	-	-	n = 125
ADI-R Social Interaction	-	-	-	-	n = 123
VABS-2 Socialisation standard score	-	-	-	-	n = 105
SRS-2 Total T score	-	-	-	-	n = 107
CBCL ADHD subscale	-	-	-	-	n = 118

3.3 Results

3.3.1 Autism and ADHD likelihood

Of the infants who provided data from the EEG task at 14 months (n=100), two infants were excluded from the analysis who turned out to be half-siblings of children with a diagnosis, resulting in a sample of n = 98 infants.

Looking at the stimulus. The proportion of looking at the stimulus (transformation: orderNorm) did not differ significantly by stimulus condition and/or autism likelihood and/or ADHD likelihood (all $p > .05$). Of note, there was a marginally significant interaction between stimulus and ADHD likelihood ($p=.075$, $F (1,94) = 3.23$, $\eta_p^2 = .03$), with infants with elevated likelihood of ADHD looking more at the non-social versus social stimulus.

Theta power. Overall relative theta power was significantly reduced in infants with elevated likelihood of ADHD ($p = .03$, $F= 4.84$, $\eta_p^2 = 0.39$). Furthermore, there was a significant interaction between stimulus condition and autism likelihood ($p = .02$, $F = 5.83$, $\eta_p^2 = .06$; main effect stimulus: $p < .0001$, $F = 24.24$, $\eta_p^2 = 0.21$). Follow-up ANOVAs by group revealed that while infants of both groups showed significantly stronger theta power while looking at the experimenter versus toy, the effect was less pronounced in infants with elevated ($p = .01$, $F = 6.40$, $\eta_p^2 = 0.09$) compared to typical likelihood of autism ($p = <.0001$, $F = 32.13$, $\eta_p^2 = 0.51$; Figure 3.1).

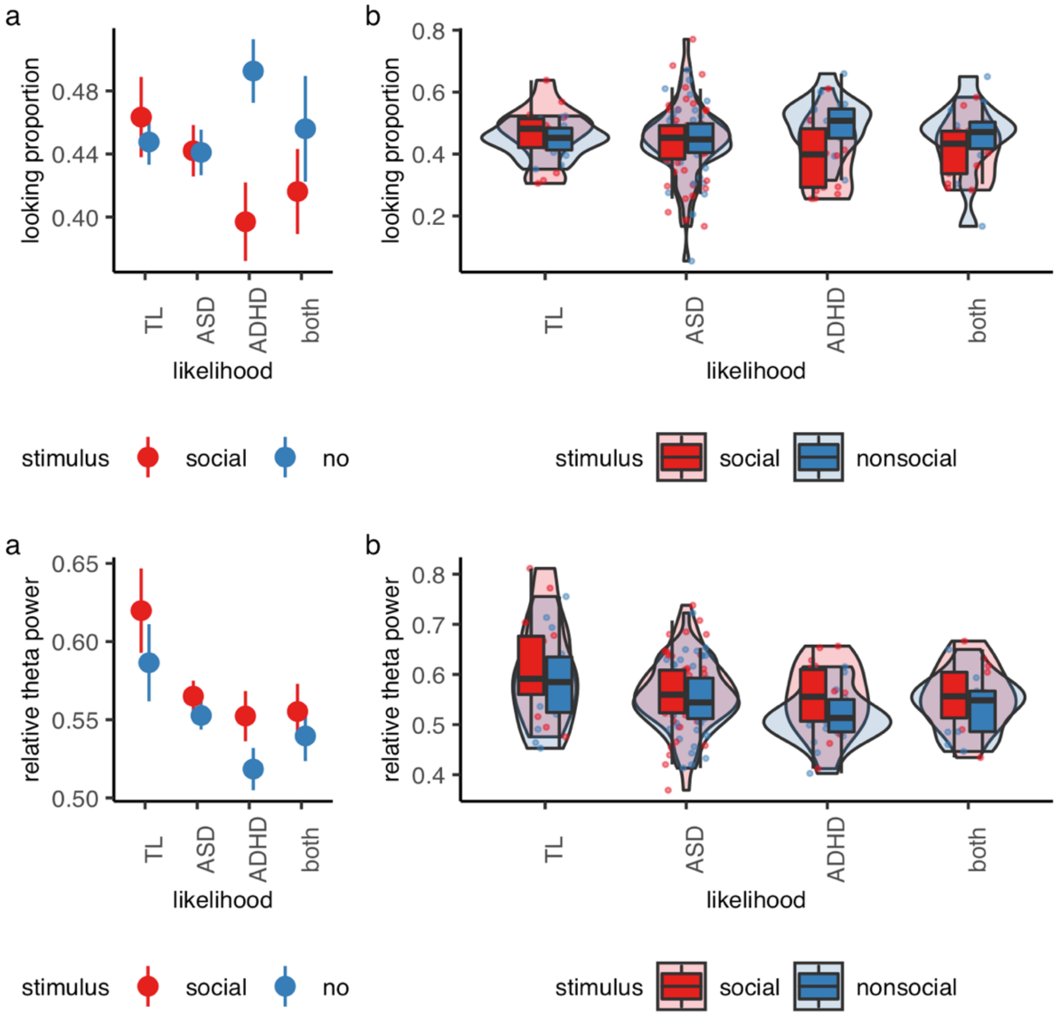


Figure 3.1. The proportion of looking at the live stimulus (top) and relative theta power during looking at the live stimulus (bottom) by stimulus condition (social: experimenter, non-social: toy) and by autism and/or ADHD familial likelihood. Left: means and standard errors, right: raincloud with boxplot (median), and data points.

Region (frontal, parietal, posterior, temporal) and side (left, central, right) as well as their interaction were added to the model to see whether these moderate the effect of reduced theta power while looking at the toy versus experimenter in infants with elevated likelihood for autism, but none of these significantly interacted with stimulus condition and autism likelihood (all $p > .3$, see A.3.1 for figures). There was a significant interaction between stimulus and region ($p < .005$, $F = 4.35$, $\eta_p^2 = .006$). Follow-up ANOVAs revealed that theta power was stronger during social versus non-social processing in all four regions, but the effect was strongest over posterior electrodes (η_p^2 = .28; frontal: η_p^2 = .09; parietal: η_p^2 = .15; temporal: η_p^2 = .10). Finally, there were significant interactions between side and region ($p < .0001$, F = 36.16, η_p^2 = .09; region: $p < .0001$, F = 137.20, η_p^2 = .17; side: $p < .0001$, F = 162.44, η_p^2 = .14), between ADHD and side ($p < .01$, F = 5.19, η_p^2 = .005; ADHD: p = .03, F = 4.84, η_p^2 = .05), and between autism, ADHD and region (p = .047, F = 2.65, η_p^2 = .004; autism likelihood X region: p= .03, F = 3.04, η_p^2 = .004).

Of note, while Levene's Test suggested that the variances of the groups with typical and elevated likelihood of autism were not homogeneous (p = .02), visual inspection indicated that the variance of the group with elevated autism likelihood was only slightly smaller than in the group with typical likelihood. One explanation could be the differing group size: the group with elevated autism likelihood included 63 infants, the group with typical autism likelihood 31 infants.

3.3.2 Autism diagnosis at 36 months (yes/no)

Information about autism diagnosis at 36 months was available for n = 128 infants of the cohort, of which 13 infants received an autism diagnosis, and 115 infants did not. For the remaining 38 infants an assessment could not take place because they did not have a visit at 36 months. Of the 98 infants who provided data from the live singing EEG task at 14 months, 85 infants had diagnosis data at 36 months. Of these, 8 infants went on to receive a diagnosis, while 77 infants did not. There was one extreme outlier with a proportional time of looking at the non-social stimulus of only 5.43%; this infant was one of the oldest infants in the group, aged 487 days, and did not go on to receive a diagnosis.

Looking at the stimulus. The proportion of looking at the stimulus (transformation: orderNorm) did not differ significantly by autism diagnosis at 36 months (all $p > .2$). There were no other significant effects.

Theta power. Relative theta power did not differ significantly by autism diagnosis at 36 months (all $p > .4$; main effect of stimulus: $p < .0005$, F = 14.94, η_p^2 = .16). However, visual inspection of the data does suggest that the difference in theta power between social and non-social attention might be somewhat smaller in the groups with than without later autism diagnosis (Figure 3.2). There were no other significant effects.

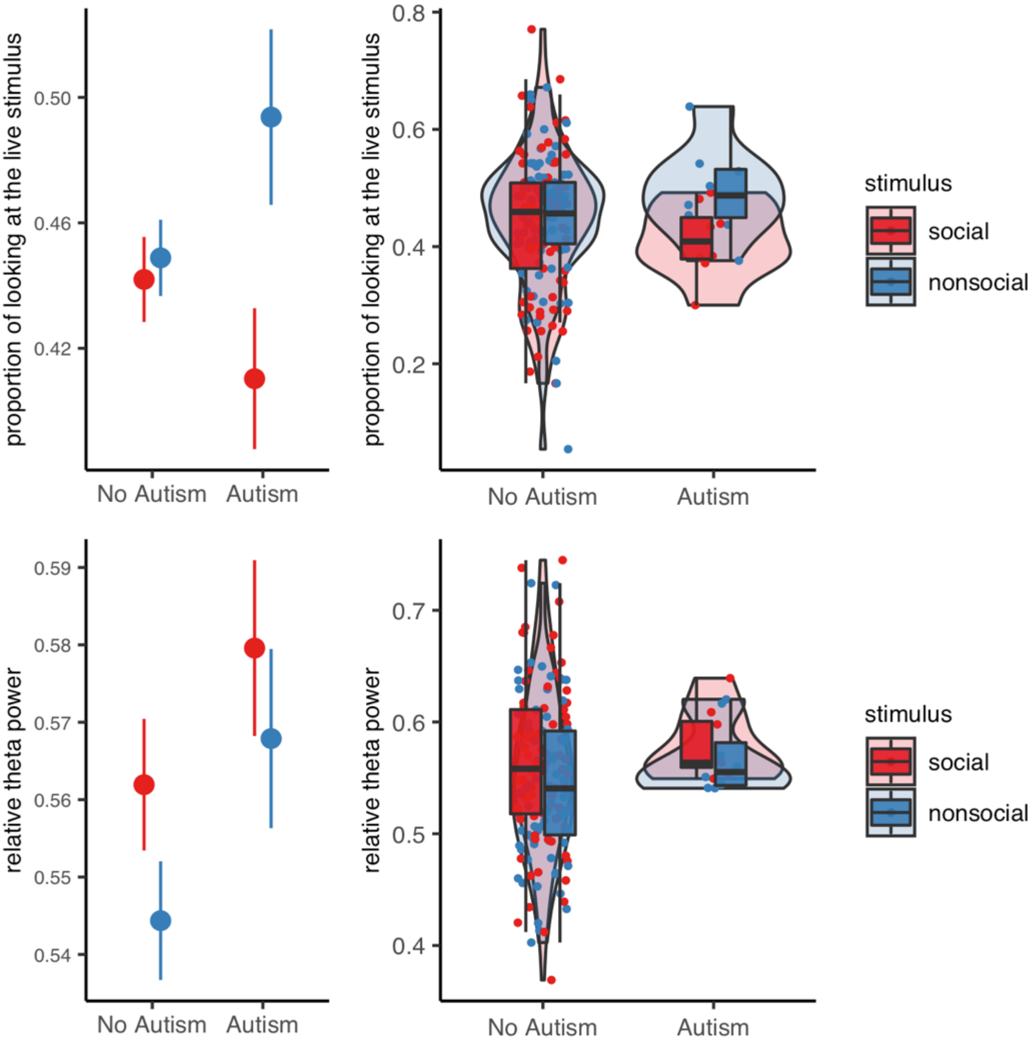


Figure 3.2. The proportion of looking at the live stimulus (top) and relative theta power during looking at the live stimulus (bottom) by stimulus condition (social: experimenter, non-social: toy) and autism diagnosis at 36 months. Left: means and standard errors, right: raincloud with boxplot (median), data points and distribution.

Given that the number of infants who went on to be diagnosed with autism (n=8) was much smaller than the number of infants who did not go on to be diagnosed with autism (n=77), normality was not given in all cells of the design, and results have to be interpreted with caution.

3.3.3 Autism traits at 36 months

Of the 100 infants who provided valid EEG data at 14 months, n = 84 infants provided data from the ADOS and the ADI measure, respectively, and n = 72 infants from the Social Responsiveness scale and the Vineland Social scale, respectively. There was an extreme outlier with a proportional time of looking at the non-social stimulus of only 5.43% this infant was one of the oldest infants in the group (487 days) and scored low on the ADOS at 36 months and very low on the ADI at 36 months.

Looking at the stimulus. There was a significant, although weak, interaction effect between stimulus condition and SRS Total score ($p = .048$, $t = 2.01$, $\eta_p^2 = .04$). Children who looked less at the experimenter and more at the toy at 14 months of age went on to score higher on the SRS, that is have more severe impairments of social responsivity, at 36 months (Figure 3.3a). No other effects were observed (all > .1).

Theta power. There was a significant interaction between stimulus condition and ADOS score at 36 months ($p = .048$, $t = 2.02$, $\eta_p^2 = .07$), with a reduced elevation of theta power during social versus non-social attention at 14 months relating to higher ADOS scores at 36 months (Figure 3.3b). Of note, when using only the ADOS score in the model, while dropping the other measures, the interaction is not significant anymore ($p>.2$), although then more infants contributed data from the ADOS variable. This suggests that the significant ADOS effect might have possibly been driven by infants who did not provide data on the other measures. The result should on any account be interpreted with caution. No further effects were observed.

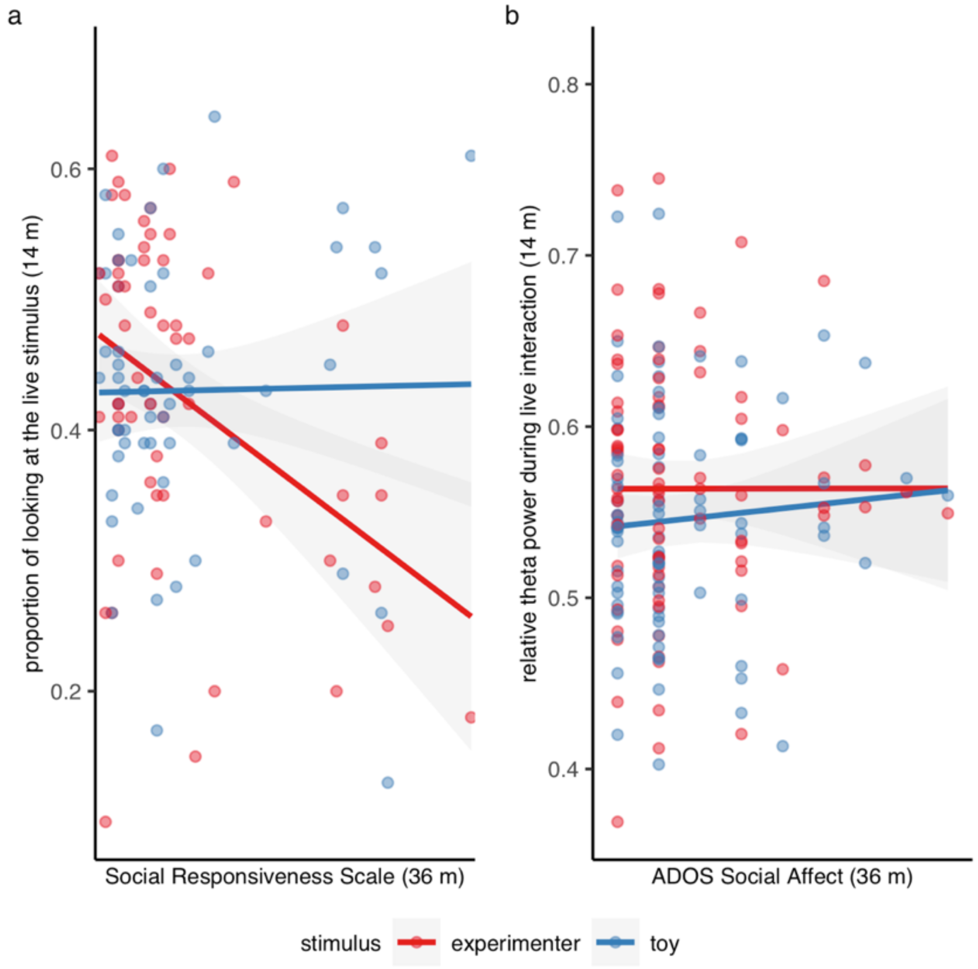


Figure 3.3. a) The proportion of looking at the stimulus by social responsiveness at 36 months. b) Relative theta power by social affect measured by the ADOS at 36 months.

3.3.4 ADHD traits at 36 months (specificity analysis)

Of the 100 infants who provided looking and EEG data in the 14-month live singing task, n=77 infants contributed data from the CBCL ADHD scale.

Looking at the stimulus. The proportion of looking at the social versus non-social stimulus during naturalistic experiences at 14 months of age was not significantly related to the CBCL ADHD score at 36 months ($p = .07$, $t = 1.80$, $\eta_p^2 = .02$).

Theta power. Theta power during social versus non-social naturalistic experiences at 14 months was not significantly related to the CBCL ADHD score at 36 months ($p = .08$, $t = 1.76$, $\eta_p^2 = .04$; main effect stimulus: $p < .0005$, $t = -3.86$, $\eta_p^2 = .17$).

3.3.5 Trajectory of looking at the face

Of the infants who took part in the eye-tracking task “50 Faces” (10m: n=115; 14m: n=124; 24m: n=91), I excluded from the analysis infants who looked at the screen for less than 5% of the time during the task. The resulting number of participants entering the analysis per timepoint was n=113 at 10 months, n = 123 at 14 months, and n = 91 at 24 months. Overall, the analysis included data from N = 148 infants who provided data for at least one of the three timepoints. The four eye-tracking measures of interest were 1) proportion of looking at the face; 2) number of looks to face; 3) mean fixation duration to the face; and 4) peak fixation duration to the face).

Peak look duration to a face, mean look duration to a face and the proportion of looking at a face, respectively, were not predicted by the proportion of looking or theta power during live action at 14 months (all $p > .2$). On the number of looks to a face, there was a marginally interaction effect between the proportion of looking at the live stimulus, stimulus condition and eye tracking timepoint ($p = .08$, $t = -1.76$, $\eta_p^2 = .01$), but following up the interactions did not reveal significant effects of looking at the live stimulus.

3.4 Discussion

3.4.1 Summary

The present study investigated EEG responses of 14-month-old infants with elevated and typical likelihood for autism and/or ADHD attending to a live experimenter who was either singing nursery rhymes (social) or operating a dynamic toy (non-social). Previous work reported increased theta power during social versus non-social live action in typically developing 12-month-olds (Jones et al., 2015) and reduced theta power modulation during social versus non-social attention in *toddlers* with autism compared to toddlers without autism (Dawson, Bernier, et al., 2012). Therefore, I predicted reduced elevation of theta power during social attention at 14 months in infants with elevated autism but not ADHD likelihood, in infants who went on to be diagnosed with autism and in infants going on to score higher on autism trait measures but not ADHD trait measures at 36 months. I also predicted that socially selective theta power strength in infancy predicted more looking at the face at different timepoints in toddlerhood.

Results showed that theta power, but not looking proportion, at 14 months was higher during attending to a live experimenter versus toy in the overall sample. Second, the stimulus effect was present in both infants with typical and elevated autism likelihood, but weaker in the latter. This alteration by social context was only observed on the neural level, while the proportion of *looking* at the experimenter versus toy did not differ between autism likelihood groups. Third, theta power in the small group of infants who went on to be *diagnosed* with autism at 36 months (n=8) did not differ significantly from infants who were not diagnosed. Fourth, infants looking less at the experimenter and more at the toy at 14 months went on to score higher on the SRS at 36 months, while theta power was not related to later SRS score. Infants with reduced theta power elevation during social versus non-social attention at 14 months scored higher on the ADOS social affect scale at 36 months. However, this effect might not be robust and has to be interpreted with caution. Neither looking nor theta power patterns during naturalistic experiences at 14 months were related to ADHD traits at 36 months. Finally, a higher proportion of looking at the experimenter versus toy at 14 months was associated with more looks to a face in a video at 14 and 24 months, and not associated with the other face looking behaviours. Theta power during live naturalistic experiences was not related to the trajectory of looking at a face.

3.4.2 Sensitivity of theta power to social information

As predicted theta power in the overall sample of 14-month-old infants was enhanced during looking at the experimenter compared to looking at a toy in a live social context. This finding is in line with the findings from Jones et al., 2015, showing enhanced theta power activation during social versus non-social live stimulation in a typical sample at 6 months and 12 months.

In the present sample, the effect was present in all four regions while appearing strongest in the posterior/occipital region of the cortex. By contrast, at 12 months, it was observed only over frontal, parietal and occipital, not temporal, and was strongest over the frontal region (Jones et al., 2015). Importantly, the effect was also shown to be more widespread at 12 compared to 6 months, suggesting that in contexts combining various multi-modal social cues, social brain activity expands over the second half of the first year of age (Jones, Venema, Lowy, Earl & Webb, 2015), and it is possible that this expansion continues after the end of the first year, resulting in the widespread differential activation observed at 14 months. This is in line with predictions of the Interactive Specialisation hypothesis, according to which over the course of infancy, neural activation increases in specificity to particular stimuli (e.g., social cues) and this specificity increases in spatial extent (Johnson et al., 2009). At 12 months, the effect of stronger theta power during social versus non-social attention was stronger over frontal electrodes (Jones, Venema, Lowy, Earl & Webb, 2015), while the present findings show that at 14 months, it appears strongest over posterior electrodes, might be explained by the slight difference between the paradigms in these studies. Besides this difference in paradigm, the current sample was more heterogeneous than in Jones et al. (2015) in that it also included infants with elevated likelihood for neurodevelopmental conditions, and it is possible that there are low-level differences between likelihood groups in the role different regions play for selective social processing.

Viewing the present findings together with a recent study analysing theta power during social versus non-social video viewing at 14 months highlights the importance of naturalistic contexts for studying social attention in infancy. The authors of that study observed elevated theta power during the social versus non-social condition, with the effect being strongest occipitally, followed by frontally, followed by parietally, and did not appear temporally (Haartsen et al., 2022). Of note, this was the same cohort as presented in this study, which means that differences between the studies should not be related to sample characteristics. Both in the screen and in the live context, the effect was consistently strongest occipitally. Viewed together with the findings of the effect being strongest over frontal regions at 12 months, this indicates an age-related change in the role different regions have for social versus non-social processing. Moreover, the video context elicited the effect in three of the four regions investigated, the live context elicited the effect in all four regions, including temporally. This adds to previous findings (Jones et al., 2015) suggesting that live stimulation may elicit more widespread effects of differential theta power, and should be strongly considered when investigating social attention development.

The amount of looking at the live stimulus did not differ between the social versus non-social condition in the present 14-month-old sample, while 6- as well as 12-month-olds in the previous sample of typical infants were looking more at the non-social versus social live stimulus, with the effect being stronger at 12 months (Jones, Venema, Lowy, Earl & Webb, 2015b). Again, this difference might be explained by differences in paradigm, age or heterogeneity of the sample. In the study with 6- and 12-month-old infants, the social/non-social conditions were derived from infants' looking episodes at either experimenter or the plain toy while the experimenter was singing throughout and looking at the infant, and infants actively had to choose looking at one or the other in order to enter the respective condition. By contrast, in the current study with 14-months olds, the experimenter was either singing and looking at the child or spinning a toy and looking at the toy. Hence, some aspects in the current non-social condition might have been less engaging than previously (no simultaneous singing versus simultaneous singing; experimenter looking at the toy versus experimenter looking at the child). The use of the dynamic versus previously used plain toy appears to not have elevated infants' visual attention. Taken together, the present results add to the finding that theta power is particularly elevated by social stimuli, even though visual attention towards social versus non-social was equal (present study) or reduced (Jones et al., 2015), and that differential theta power appears as a more insightful measure than looking behaviour to study attention engagement in a naturalistic social context. It remains unclear what aspects within naturalistic social contexts drive infants' elevated attention in these contexts, and it is important to systematically investigate the effect of various aspects.

A theoretical account to address this question of why infants are more engaged by social contexts is provided by the theory of infant natural pedagogy. This view proposes that social cues are ostensive, signalling to the child that they are being addressed and can expect information, and that infants inherently attend to communicative/ostensive cues that evoke in them a feeling of being addressed (Csibra and Gergely, 2009). During naturalistic social interaction, the infant perceives multiple multimodal communicative cues at the same time which might trigger a particularly strong feeling of being addressed and hence a heightened state of attention, reflected by higher theta power. As such, theta power might be stronger during social versus non-social stimulation because it involves a myriad of cues signalling to the infants that they are being addressed and communicated to.

3.4.3 Differences by familial autism/ADHD status

This is the first study investigating theta power during naturalistic social and non-social experiences in infants with elevated autism likelihood. Results showed that the effect of elevated theta power during social versus non-social attention was reduced in infants with familial likelihood of autism. This finding is consistent with previous research in *toddlers* with autism diagnosis, who showed a reduced increase in theta power to social versus non-social stimuli (Dawson, Bernier, et al., 2012). Further, receiving a longer-term behavioural intervention led to improved social behaviour along with more typical patterns of theta power during social versus non-social stimulation in these children (Dawson et al., 2012).

Further, the present findings are in line with findings reporting that infants with elevated autism likelihood who received an early parent-delivered intervention targeting parent-child interaction showed a greater increase of theta power during social versus non-social videos from 6 to 12 months compared to infants receiving a control intervention, as well as concurrent improvement of other neurocognitive measures of social attention (Jones et al., 2017), suggesting that increased differential theta power correlates with improved socio-cognitive skills and might be causally related to social experiences. Finally, the present result is in line with research using other techniques to study alterations in social brain function in infants who have an elevated likelihood for autism (Lloyd-Fox et al., 2013) and go on to be diagnosed with autism in toddlerhood (Jones et al., 2016; Lloyd-Fox et al., 2018). These infants showed reduced elevation of neural attention engagement in social versus non-social stimuli, reflected by fNIRS (Lloyd-Fox et al., 2013; 2018) and ERP measures (Jones et al., 2016). Taken together, theta power in social versus non-social live contexts is a promising candidate in the search for measures detecting alterations in typical social brain processing in infancy.

Crucially, the present findings of reduced theta power elevation during social versus non-social live processing in infants with elevated likelihood for autism at 14 months are specific to the live format. In fact, a recent study with the same cohort at the same timepoint, viewing social and non-social videos, did not reveal reduced theta power during social versus non-social video viewing in infants with elevated likelihood for autism (Haartsen et al., 2022)

This finding adds to the previous finding in typical infants, where live but not screen-mediated stimulation has shown to elicit elevated theta power during social versus non-social stimuli at 6 months (Jones et al., 2015). Together, these findings highlight the importance of studying social brain function in a naturalistic context, particularly for studying early alterations in social processing (Jones et al., 2015).

The reduced effect between social and non-social processing in infants with elevated likelihood for autism in the present study was specific to theta power, while the proportion of looking at the experimenter versus toy did not significantly differ across groups. This is in line with previous research indicating that differences in social processing on the neural level before they manifest in visible behaviour (e.g. Jones et al., 2015), and that neural measures might be more sensitive in measuring infants' attention than looking behaviour alone (Begus & Bonawitz, 2020; Begus et al., 2016).

Further, the reduced effect on the level of theta power was specific to infants with elevated autism likelihood, while it was not reduced in infants with familial likelihood of ADHD. This finding suggests that the alterations in socially selective theta power are specific to early autism, and not shared among neurodevelopmental conditions. I did, however, observe reduced *overall* relative theta power in infants with elevated likelihood of ADHD. This finding is in line with recent research reporting a lower theta-beta power ratio in 10-month-old infants with familial likelihood of ADHD, which in turn was related to ADHD traits at 24 months (Begum-Ali et al., 2022). It has been proposed that while impairment of social attention is specific to autism, ADHD symptoms might be the result of alterations in the domain-general attention system (e.g., Braithwaite, Gui & Jones, 2020).

The current findings in 14-month-old infants are a further step towards identifying early biomarkers of altered social development. Theta power during naturalistic social versus non-social experiences in infancy is a promising measure to capture individual differences in the engagement with a live social partner at 14 months. However, in typical infants elevated theta power during social versus non-social naturalistic experiences has been observed at the age 6 months of already (Jones et al., 2015), and studies using other techniques have observed alterations in social versus non-social processing in infants with elevated likelihood for autism or later diagnosis around this age (Jones et al., 2016; Lloyd-Fox et al., 2013, 2018). Therefore, it is possible that differences in theta power between social and non-social live contexts are altered in infants with elevated autism likelihood already at 6 months. An important next step is hence to analyse theta power in live social versus non-social situations in infants with typical and elevated likelihood for autism at the age of 6 months.

3.4.4 Differences by autism and ADHD outcome

The present study did not reveal significantly reduced theta power in infants who went on to receive an autism diagnosis at 36 months. However, the group size of these infants was very small (n=8), making it hard to reveal potential effects. In fact, visual inspection of the data showed a reduced theta power elevation during social versus social attention in the infants who went on to be diagnosed. This descriptive pattern is in line with the findings on familial likelihood (see above) and research reporting altered social processing in toddlers with autism diagnosis (Dawson, Bernier, et al., 2012). Further studies are needed to test the effect in a sample including more diagnosed infants, providing more statistical power to reveal a potentially small effect.

Besides categorical autism outcome, population-wide traits were studied which altered in autistic toddlers. Infants who looked less at the experimenter and more at the toy at 14 months of age went on to show more severe social responsiveness impairments at 36 months as measured by the SRS. Interestingly, infant theta power differences were not related to the SRS score at 36 months. Further, infants who showed a reduced elevation of theta power during social versus non-social naturalistic experiences at 14 months went on to show more impairments of social affect at 36 months as measured by the ADOS.

ADHD outcome was not included in the present analyses since a diagnosis is usually given only after 4 years of age (Wolraich et al., 2011). ADHD traits measured at 36 months were not related to infant selective looking or theta power during social versus non-social live action. This supports views suggesting that altered social processing in infancy is specific to early autism instead of a characteristic of a shared phenomenon across neurodevelopmental conditions.

3.4.5 Limitations and future directions

The present study did not reveal significant effects in relation to autism outcome. This was likely because the size of that group was very small. Future studies are should to investigate theta power responses at 14 months in a group with more infants who go on to be diagnosed.

The findings regarding the relation with outcome trait measures have to be interpreted with caution, since the effects were very small and did not seem robust. In fact, when ADOS was only measure in the model, the effect vanished, suggesting that few infants were driving the effect and it might not be reliable.
Studying brain function to combined multimodal stimuli in naturalistic contexts has the obvious advantage of high ecological validity of the results. On the contrary, while the present study reveals effects in respect to social versus non-social processing in general, the question remains about what aspect in social interaction is driving this effect.

Finally, an important next step will be to analyse theta power in live social versus non-social situations in infants with typical and elevated likelihood for autism at an earlier age, to see when these differences first emerge. Knowing whether these differences are present already in early infancy is important in order to be able to identify alterations and provide potential support early on.

3.4.6 Conclusion

The present study investigated theta power activity in a live social versus non-social context in 14-month-old infants with and without elevated likelihood for autism and/or ADHD. The results replicate previous findings showing elevated theta power during naturalistic social versus non-social experiences in the overall sample, and further suggest that differential theta power to social versus non-social live stimulation is reduced in infants with elevated compared to typical likelihood for autism. A next step is to investigate when these differences first emerge in infancy, and whether they are linked to autism diagnosis in childhood.

CHAPTER 4: THETA POWER RESPONSES DURING NATURALISTIC SOCIAL EXPERIENCES IN INFANTS WITH TYPICAL AND ELEVATED LIKELIHOOD FOR AUTISM AT 5 MONTHS

4.1 Introduction

When searching for reliable measures to reflect early attention engagement in social versus non-social stimuli, relative theta power during social versus non-social naturalistic experiences could be a promising measure, as it has shown to differ between infants with typical versus elevated likelihood for autism at 14 months of age (Chapter 3). Because differences in how social information is being processed likely emerge already during the first year of age, measures are needed to study differences in attention engagement at that age. In the case of atypical processing, interventions can be most effective when implemented at a stage where differences are still small and there is still great plasticity (Jones et al., 2017).

4.1.1 Sensitivity of theta power to social cues increasing over the second half of the first year

There is great developmental change in the time window between 5 and 12 months. Already at a few months of age, infants display preferences and skills related to relevant social stimuli; for example, being able to recognise individual novel faces (De Haan, Johnson, Maurer & Perrett, 2001). In the second half of the first year, infants start to interact more with the world around them. They learn the basics of social interaction, such as joint attention and turn taking, and these behavioural changes are accompanied by dramatic changes in the brain. One of these changes relate to theta power. For example, between 5 and 10 months, the theta network increases in connectivity and its sensitivity to social versus non-social processing (Van der Velde et al., 2021). Further, theta power during attention to a social partner versus toy increases from 6 to 12 months in typically developing infants (Jones et al., 2015). These findings indicate an increase in social sensitivity and spatial extent of the infant social brain network in the second half of the first year of age.

4.1.2 Reduced theta power modulation indicating social attention alterations at 6 months?

Previous findings suggest that alterations in neural processing of social versus non-social stimulation are present already at the start of the second half of the first year. For example, infants with later autism diagnosis showed altered ERP responses to eye gaze shifts to versus away from the infant at 6-10 months (Elsabbagh et al., 2012) and at 8 months (Tye et al., 2020), engaged less with a face stimulus at 6 months (Jones et al., 2016), and showed reduced responses to social versus nonvocal sounds at 4-6 months (Lloyd-Fox et al., 2018). Together, findings of increasing social sensitivity of theta power between 6 and 12 months and processing differences being present already present at 6 months suggest that theta power during social versus non-social attention might be altered in infants with emerging social attention difficulties as early as at 6 months.

4.1.3 The present study

The present study analysed whether the reduction in theta power while looking at an experimenter versus toy in 14 months old infants with elevated likelihood for developing autism is present as early as at 4-6 months. As reported in Chapter 3, in a within-subject design infants were presented with an experimenter consecutively either singing nursery rhymes (social) and operating a dynamic toy (non-social) while EEG responses were recorded.

The live context was used to increase ecological validity and because it has proven to be more powerful in eliciting differential responses to social versus non-social action (Jones et al., 2015). Infants at 5-6 months of age were recruited. Sensitivity of theta power to social processing has previously been shown to be present at 6 months of age in typical infants (Jones et al., 2015). Examining slightly younger infants will further narrow down the timescale over which the enhanced social response in the theta band emerges. Further, atypicalities in social versus non-social responses have been shown around six (Jones et al., 2016) and at 4-6 months (Lloyd-Fox et al., 2018) in infants with later autism diagnosis, and at 4-6 months in infants at elevated likelihood for autism (Lloyd-Fox et al., 2013). Therefore, if frontal theta power is a reliable marker of social processing, differences are expected to reveal at that age.

I predicted that theta power in a social versus non-social live context 1) is reduced in infants at elevated likelihood for developing autism, but not ADHD; 2) is reduced in infants who go on to develop autism or score high on autism but not ADHD trait measures at 36 months, and 3) predicts the trajectory of looking at a face at 10, 14 and 24 months.

4.2 Methods

4.2.1 Participants

As in Chapter 3, infants were part of the *Phase-3/STAARS* (*Studying Autism and ADHD in the eaRly yearS*) cohort of the longitudinal *British Autism Study of Infant Siblings* (*BASIS*). The current analysis focused on EEG data obtained at the 4-6-month timepoint. The current chapter focusses on the 5 months timepoint. In total, 92 infants contributed data from the live singing EEG task at this visit. The remaining 74 infants did not have data because they did not have this timepoint visit (n = 58), did not tolerate net or baby was upset (n = 9), technical fault related to the acquisition software (n = 2), because there was no video (n = 1) or because the EEG file was corrupted (n = 4).

4.2.2 Autism and ADHD likelihood assessment

The sample of 92 infants with valid EEG data at this timepoint included infants at typical autism likelihood (n = 20), elevated autism likelihood (n = 43), elevated ADHD likelihood (n = 15) or elevated likelihood for both conditions (n = 12). Of note, as in Chapter 3, analyses were run with factors of autism presence or absence as well as ADHD presence or absence. *See Chapter 3 for a detailed description of autism and ADHD likelihood assessment*.

4.2.3 Procedure and measures

The measures that were analysed in relation to EEG responses in the live singing task at 4-6 months were the same as at the 14-month timepoint; *see Chapter 3*. Infant numbers by outcome measures and per likelihood group are listed in *Table 4.1a* and 4.1*b*, respectively.

4.2.3.1 EEG pre-processing and analysis

Processing and analysis steps were the same as for the EEG data obtained at the 14-month timepoint; see Chapter 3. Infant numbers per likelihood group and by outcome measures are listed in Table 4.1a and 4.1b, respectively.

Table 4.1a. Number of infants with valid EEG data at 5 months, by likelihood group; *two additional infants with valid EEG data at 5 months were excluded because they were half-siblings or had unclear likelihood status

Timepoint					
Likelihood group	Total	TL	EL-ASD	EL-ADHD	EL-both
Available EEG data at 5 mo.	n = 90*	20	43	15	12

Table 4.1b. Number of infants providing data on the tasks and measures used. ADOS-2 = Autism Diagnostic Observation Schedule; ADI = Autism Diagnostic Interview; VABS-2 = Vineland Adaptive Behaviour Scales; SRS= Social Responsiveness Scale; CBCL = Child Behaviour Checklist

	Timepoint				
	5 months	10 months	14 months	24 months	36 months
EEG Live task	n = 92		(n = 100)	-	-
Eye tracking “50 Faces” task	-	n = 115	n = 124	n = 91	-
Diagnostic assessment	-	-	-	-	n = 128
ADOS-2 Social Affect (CSS)	-	-	-	-	n = 125
ADI-R Social Interaction	-	-	-	-	n = 123
VABS-2 Socialisation standard score	-	-	-	-	n = 105
SRS-2 Total T score	-	-	-	-	n = 107
CBCL ADHD subscale	-	-	-	-	n = 118

4.3 Results

Results are reported separately for the dependent variables of a) the proportion of looking at the live stimulus, and b) theta power.

4.3.1 Autism and ADHD likelihood

From the infants who provided data from the EEG task at 4-6 months (n=92), two infants were excluded from the analysis because they were half-siblings of children with a diagnosis, resulting in a sample of n = 90 infants.

Looking at the stimulus. There was a significant main effect on the proportion of looking at the stimulus (transformation: yeo_johnson) by autism likelihood ($p = .03$, $F(1,84) = 4.83$, $\eta_p^2 = .05$), infants at elevated compared to typical likelihood for autism looking less at the stimulus, which seemed mainly driven by a reduction in looking at the social stimulus compared to infants with typical likelihood (Figure 4.1, top), although this interaction with stimulus was not significant. The interaction between autism likelihood and stimulus condition, as well as between autism likelihood, ADHD likelihood and stimulus condition, was not significant (both $p > .3$).

Theta power. There was a significant main effect on relative theta power by stimulus condition ($p < .001$, $F = 11.84$, $\eta_p^2 = .12$), with relative theta power being higher during the social versus non-social condition (Figure 4.1, bottom). The interaction between autism likelihood and stimulus condition, as well as between autism likelihood, ADHD likelihood and stimulus condition, was not significant (both $p > .1$). Interestingly, there was a significant interaction effect of autism likelihood and ADHD likelihood ($p = .03$, $F = 4.60$, $\eta_p^2 = .05$), with overall theta power being reduced in infants with elevated ADHD likelihood only if they are not also at an elevated likelihood for autism.

Of note, while Levene's test suggested that the variances of theta power between were not homogeneous between the groups of infants with typical and elevated autism likelihood ($p = .04$), visual inspection indicated that the variance of the group with elevated autism likelihood was only slightly smaller than in the group with typical likelihood. One explanation could be the differing group numbers: the group with elevated autism likelihood included 42 infants, while the group with typical autism likelihood included 20 infants.

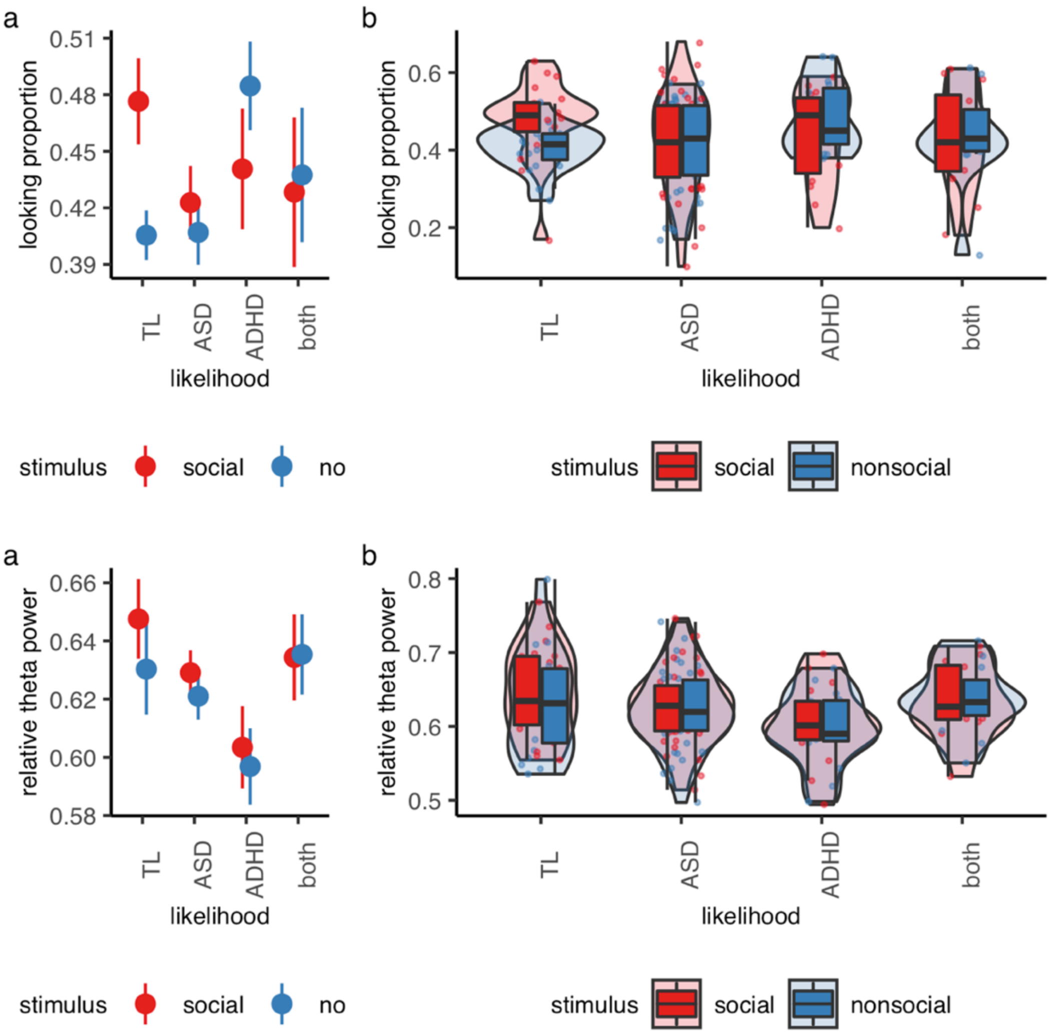


Figure 4.1. The proportion of looking at the live stimulus (top) and relative theta power during looking at the live stimulus (bottom) at 5 months by stimulus condition (social: experimenter, non-social: toy) and by autism and/or ADHD familial likelihood. Left: means and standard errors, right: raincloud with boxplot (median), data points and distribution.

Including region and side in the model revealed a significant 3-way-interaction between stimulus, region and side (Stimulus x Region x Side: $p = .03$, $F(6, 516) = 2.3$, $\eta_p^2 = .3$; Stimulus x Region: $p < .0005$, $F(3, 258) = 6.22$, $\eta_p^2 = .12$; Stimulus: $p<.001$, $F(1, 86) = 11.84$, $\eta_p^2 = .12$;autismlikelihood x ADHD likelihood: $p=.03$, $F(1, 83) = 4.6$, $\eta_p^2 = .5$; see A.4.1 for figures). Follow-up ANOVAs of the 3-way-interaction between stimulus, region and side revealed that the overall stimulus effect was present over frontal central ($p< .005$, $F(1,89)=10.2$, $\eta_p^2=.10$), and very strong over posterior left ($p<.0001$, $F(1,89)=49.89$, $\eta_p^2=.36$), central ($p<.0001$, $F(1,89)=51.17$, $\eta_p^2=.37$) and right ($p<.0001$, $F(1,89)=56.19$, $\eta_p^2=.39$), respectively.

There were no interaction effects between stimulus and autism likelihood (all $p > .2$).

4.3.2 Autism diagnosis at 36 months (yes/no)

Information about autism diagnosis at 36 months was available for $n = 125$ infants of the cohort. Of the $n = 73$ infants who also provided data from the live singing EEG task at 5 months, $n=63$ infants did not go on to receive an autism diagnosis at 3 years, $n=10$ infants did go on to receive a diagnosis. Two of the infants who did not go on to be diagnosed only provided looking data but no EEG data.

Looking at the stimulus. There was a significant effect of later autism diagnosis ($p < .005$, $F = 9.06$, $\eta_p^2 = .12$), with less looking at the stimulus in infants who go on to be diagnosed with autism (Figure 4.2, top).

Theta power. There was a significant effect of stimulus ($p = .018$, $F = 5.92$, $\eta_p^2 = .08$) with stronger theta power during social versus non-social processing, and a marginally significant effect of autism diagnosis ($p = .056$, $F = 3.80$, $\eta_p^2 = .05$) with stronger overall theta power in infants who go on to be diagnosed (Figure 4.2, bottom). The number of valid trials was significantly lower in the group with later diagnosis ($p < .0005$, $F = 12.71$, $\eta_p^2 = .08$), violating the ANCOVA assumption of independence between covariate and outcome variable. However, re-running the model once with and once without trial number as covariate did not affect results.

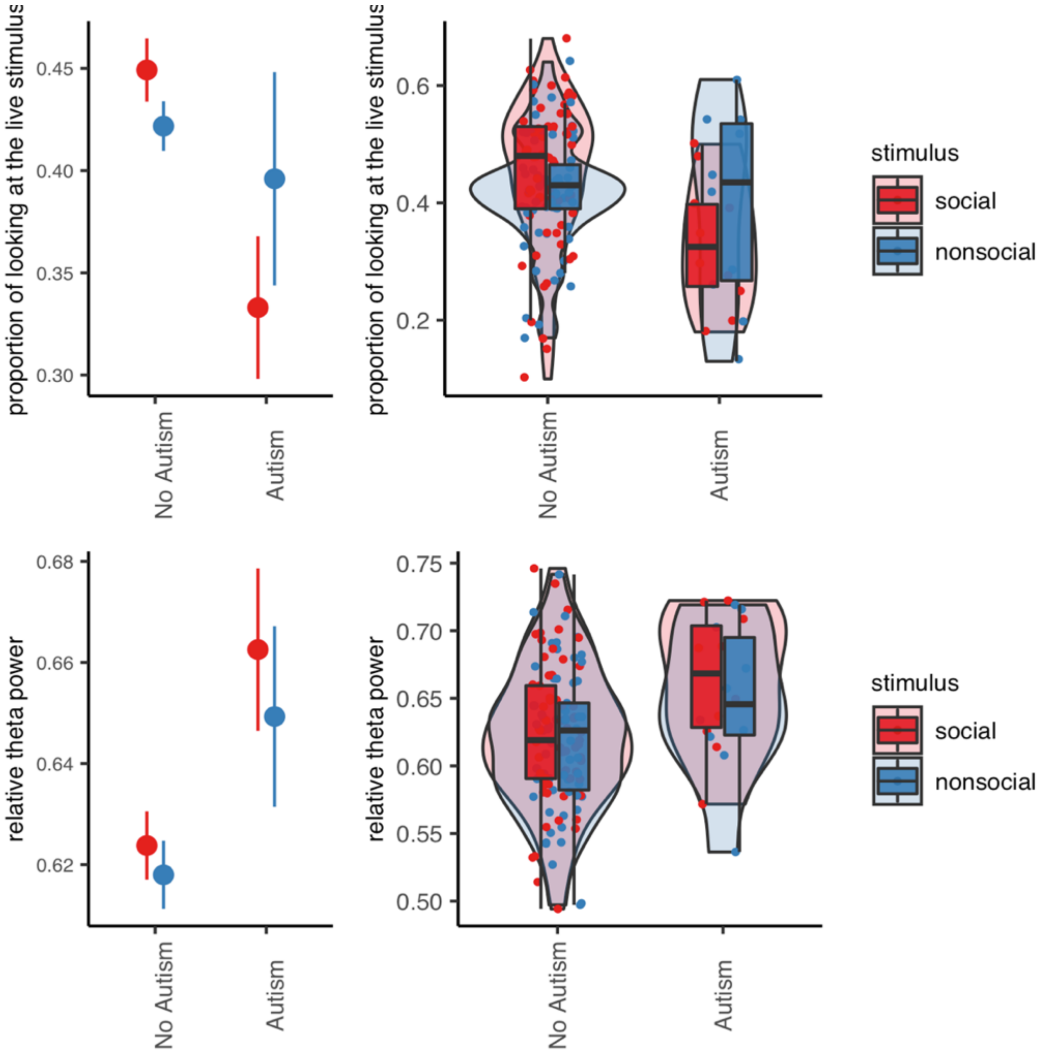


Figure 3.2. The proportion of looking at the live stimulus (top) and relative theta power during looking at the live stimulus (bottom) at 5 months by stimulus condition (social: experimenter, non-social: toy) and autism diagnosis at 36 months. Left: means and standard errors, right: raincloud with boxplot (median), data points and distribution.

4.3.3 Autism traits at 36 months

Of the n = 92 infants who provided data from the live singing EEG task at 4-6 months , n = 71 infants also provided data from the ADOS and the ADI measure at 36 months, n = 58 infants from the Social Responsiveness scale and n = 59 from the Vineland Social scale at 36 months, respectively.

Looking at the stimulus. No significant effects were observed. All interactions between stimulus condition and autism outcome measures were p > .1 (Figure 4.3, left).

Theta power. No significant effects were observed. All interactions between stimulus condition and autism outcome measures were p > .5 (Figure 4.3, right). Adding side and region did not reveal interaction effects between stimulus condition and any of the autism trait measures (all p > .4).

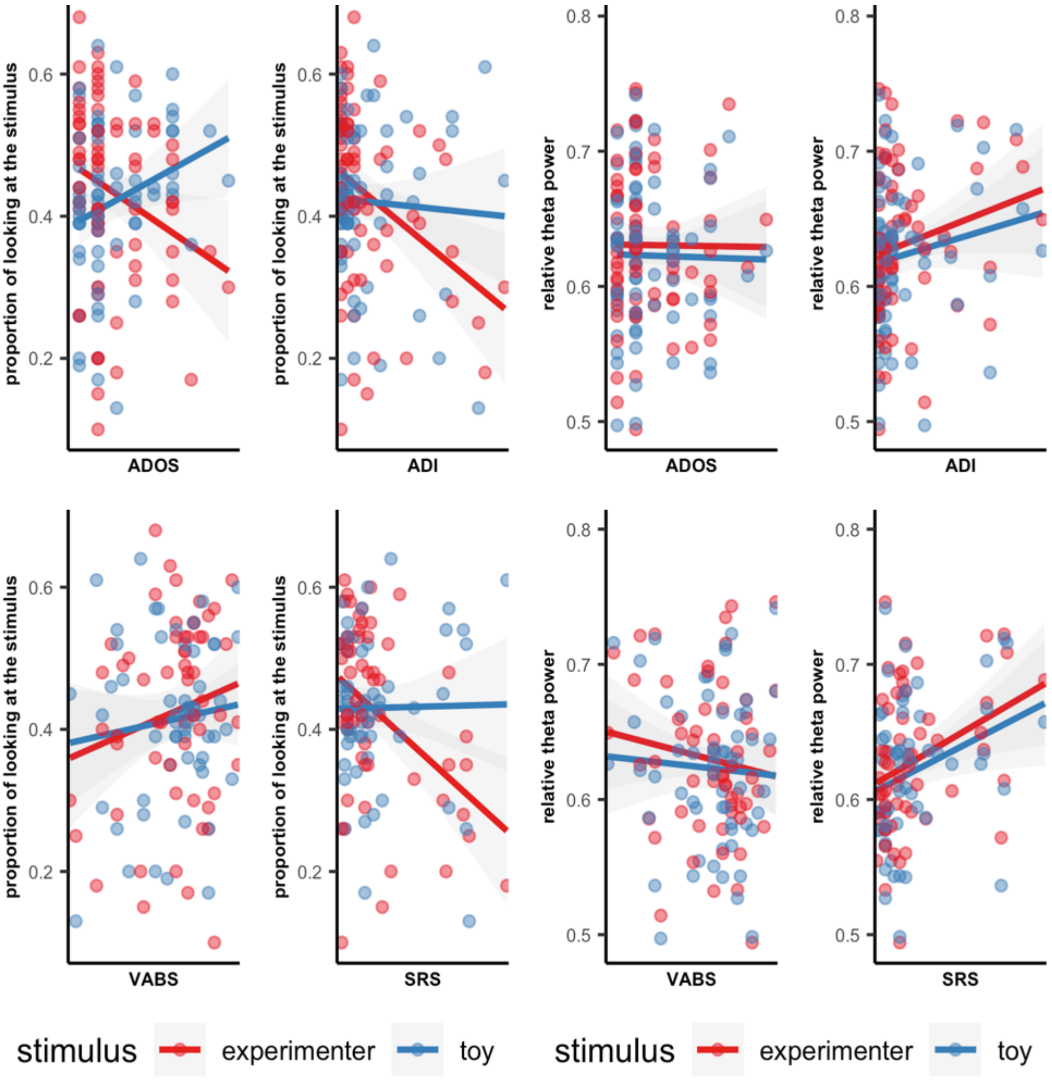


Figure 4.3. The proportion of looking at the stimulus (left) and relative theta power (right) by stimulus condition and outcome measures at 36 months.

4.3.4 ADHD traits at 36 months

Of the n = 92 infants who provided data from the live singing EEG task at 4-6 months, n = 66 infants also provided data from the CBCL at 36 months.

Looking at the stimulus. There was a main effect of CBCL ADHD score (p = .03, t = -2.2, η_p^2 = .02), with less looking at the live stimulus at 4-6 months being related to higher scores on the CBCL ADHD scale at 36 months. No further significant effects were observed.

Theta power. There were no significant effects in relation to later CBCL ADHD score on relative theta power at 4-6 months (all $p > .2$).

4.3.5 Trajectory of looking at the face

Of the infants who took part in the eye-tracking task “50 Faces” (10m: n=115; 14m: n=124; 24m: n=91), I excluded from the analysis infants who looked at the screen for less than 5% of the time during the task. Afterwards, the number of participants contributing eye tracking data to the current analysis was n=113 at 10 months, n = 123 at 14 months, and n = 91 at 24 months. The four eye-tracking measures of interest were 1) proportion of looking at the face; 2) number of looks to face; 3) mean fixation duration to the face; and 4) peak fixation duration to the face).

The peak look duration to a face, the mean look duration to a face, as well as the proportion of looking at a face, were not related to the proportion of looking at the live stimulus or theta power at 5 months (all $p > .2$). The number of looks to a face differed by an interaction of the proportion of looking at the live stimulus, stimulus condition and eye tracking timepoint ($p < .01$, $t = 2.82$, $\eta_p^2 = .03$; looking at live stimulus proportion X eye tracking timepoint: $p < .01$, $t = -2.77$, $\eta_p^2 = .0003$), and by an interaction of theta power and eye tracking timepoint ($p = .01$, $t = 2.45$, $\eta_p^2 = .04$; theta power: $p < .01$, $p = -2.65$, $\eta_p^2 = .03$). More looking at the experimenter (but not toy) at 5 months was related to more looks to a face at 10 months ($p = .01$, $F = 6.42$, $\eta_p^2 = .10$) and at 24 months ($p < .01$, $F = 8.07$, $\eta_p^2 = .15$).

Greater overall theta power at 5 months was related to fewer looks to a face at 10 months ($p < .0005$, $F= 14.55$, $\eta_p^2 = .10$). Figure 4.4 summarises face looking measures at the different timepoint by the proportion of looking at the stimulus (4.4a) and theta power (4.4b).

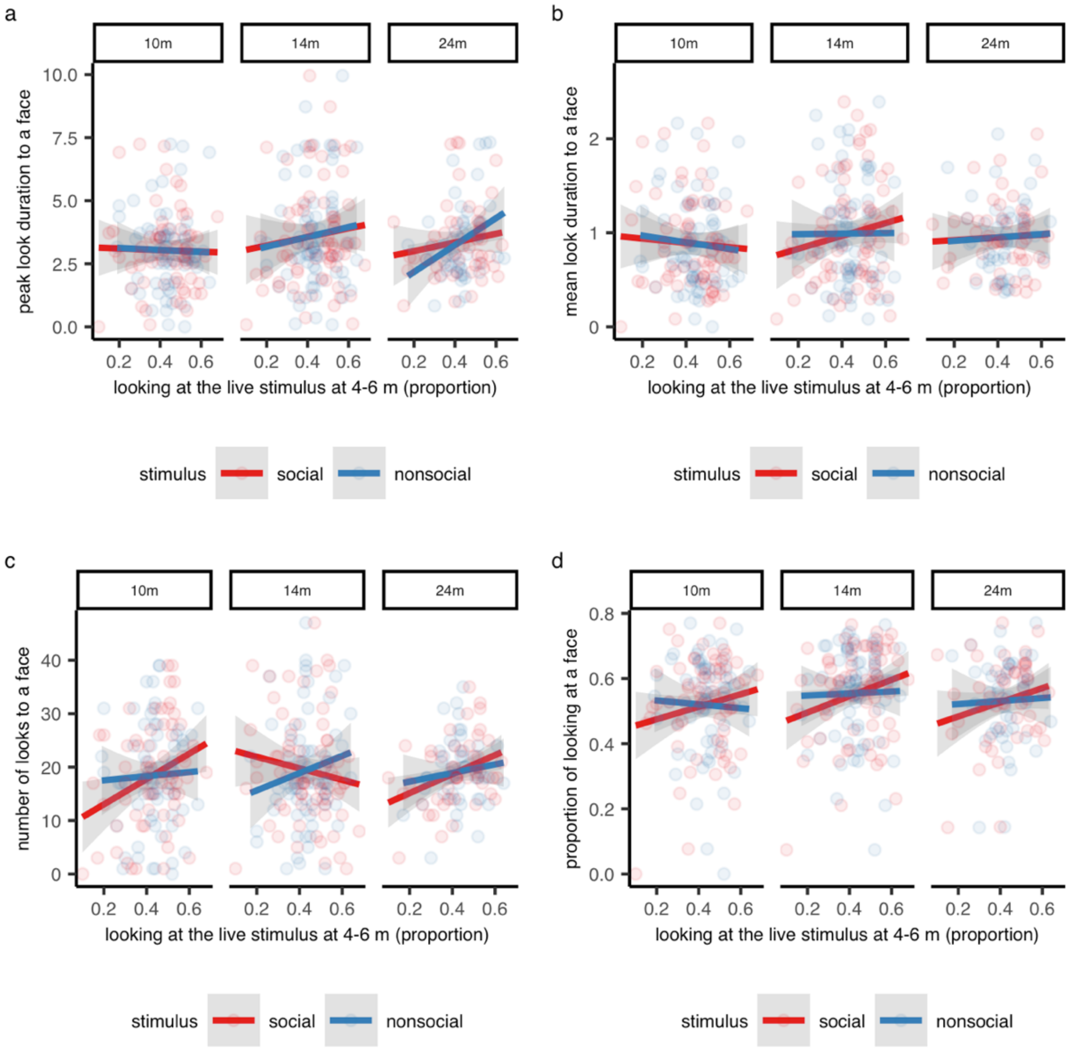


Figure 4.4a. Face looking measures at 10, 14 and 24 months, by the proportion of looking at the social versus non-social live stimulus at 4-6 months.

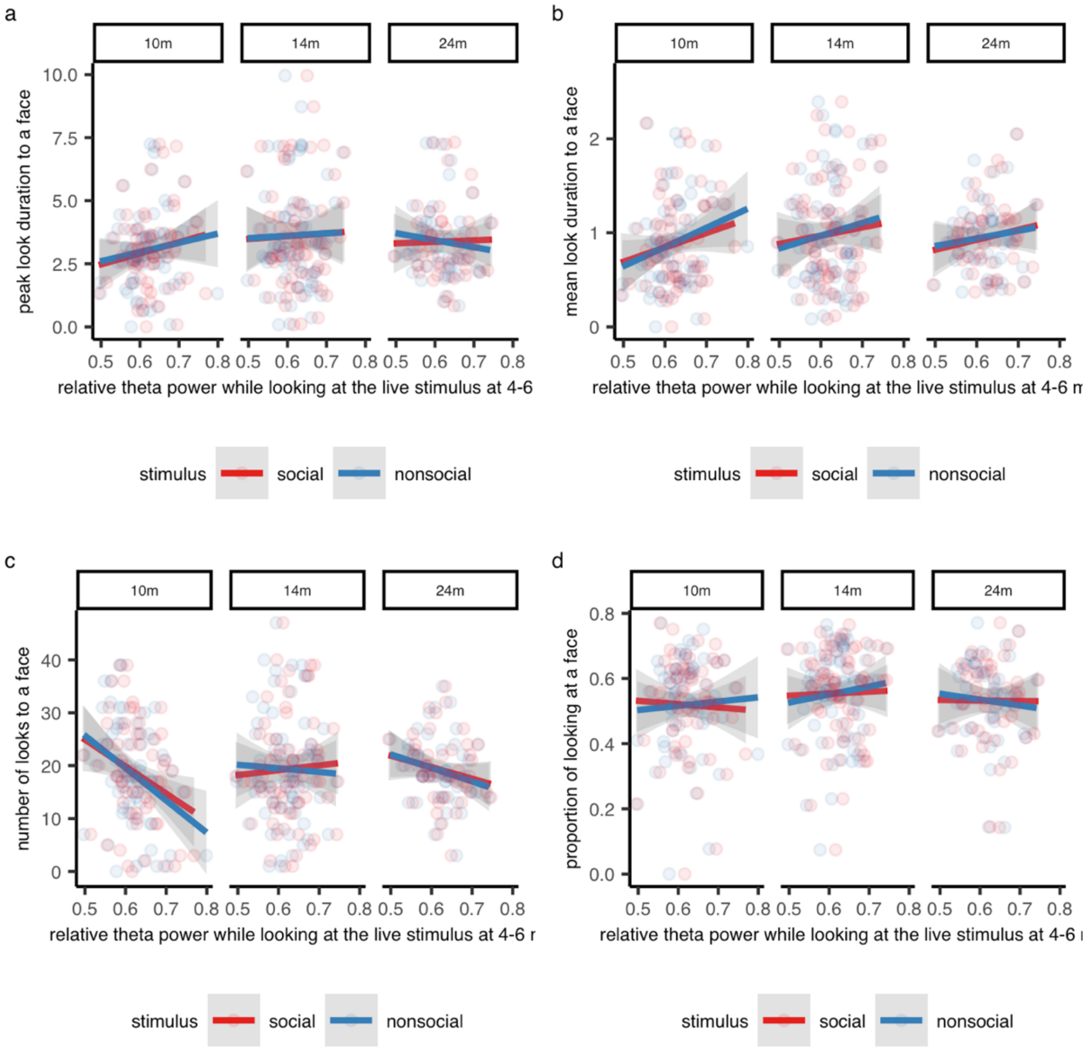


Figure 4.4b. Face looking measures at 10, 14 and 24 months, by relative theta power during looking at the social versus non-social live stimulus at 4-6 months.

4.4 Discussion

The present study investigated neural responses of 4- to 6-month-old infants with and without first-degree relative with clinical autism and/or ADHD diagnosis while attending to a live experimenter singing nursery rhymes (social) or to a live operated dynamic toy (non-social) while EEG was being recorded. Previous literature reported increased theta power during social versus non-social live action in 6-month-olds (Jones et al., 2015), a strengthening of this difference between 6 and 12 months (Jones et al., 2015), evidence of reduced theta power modulation during social versus non-social attention in toddlers with autism compared to toddlers without autism (Dawson, Bernier, et al., 2012) and infants with later autism diagnosis showing altered neural responses to social stimuli measured with fNIRS at 5 (Lloyd-Fox et al., 2018) and using ERPs at 6 months of age (Jones et al., 2016). In addition, the effect of elevated theta power during social versus non-social live action was significantly reduced in infants at elevated likelihood for autism at 14 months (Chapter 3). Based on these present and previous findings, I predicted reduced theta power during social versus non-social attention in infancy to be related to autism but not ADHD likelihood and to autism diagnosis and autism-related traits but not ADHD related traits at 3 years. Further, I predicted that theta power during social versus non-social live action would predict more looking at a face in later infancy and toddlerhood.

To sum up the present results, relative theta power at 4-6 months, but not looking proportion, was higher during the social versus non-social condition in the overall sample. The effect was present across posterior regions and sides and over frontocentral electrodes. Second, infants at elevated likelihood for autism did not show a significant alteration of this effect. However, they did show reduced overall looking at the stimulus. Infants at elevated likelihood for ADHD showed reduced overall theta power, but only if they were at typical likelihood for autism. Third, infants who went on to be *diagnosed* with autism showed less overall looking at the stimulus, too, as well as (with a marginal effect) stronger overall theta power. Fourth, autism trait measures at 36 months were not related to looking and theta power during social and non-social naturalistic experiences at 5 months, while ADHD scores were higher with reduced overall looking at the live stimulus. Finally, increased looking at the experimenter (but not toy) at 5 months was related to more looks to a face at 10 months and at 24 months. Greater overall theta power at 5 months was related to *fewer* looks to a face at 10 months.

4.4.1 Modulation of theta by social context

Theta power was stronger during social versus non-social live attention over posterior and frontocentral regions in the 6-month-old infants of the present study, replicating previous research reporting the effect over frontal regions in 6-month-old infants in a very similar paradigm (Jones, Venema, Lowy, Earl & Webb, 2015). Across the second half of the first year of age, the effect has been shown to increase both in strength and topographical extent, and by 12 months, the effect has expanded across frontal, parietal and occipital regions. By 14 months, in the present cohort revealed that the effect has expanded across all cortical regions and was then most pronounced over posterior regions (Chapter 3), in line with findings from a screen-based paradigm reporting the effect being present occipitally, frontally and parietally, but strongest occipitally at that age. The present findings hence support the view that neural responses to social stimuli become more specialised over the second half of the first year of age (Grossmann & Johnson, 2007; Johnson, Grossmann & Kadosh, 2009; Jones et al., 2015).

A difference between the present and the previous study with typical infants at 6 months is that in the present work the effect appeared fronto-centrally but strongest occipitally, while previously it was observed over frontal regions only. One reason to explain this differential occipital activation observed in the present but not in the former study could be the more heterogeneous sample including infants with elevated likelihood for autism. In line with this explanation is that in older infants differential theta power was observed to be strongest over occipital regions only in the cohorts including infants with elevated likelihood for autism (Chapter 3; Haartsen et al., 2022) while in a more homogeneous sample, older infants continued to show the strongest effect over frontal regions. This would imply that there are inter-individual differences in the extent to which different regions are involved in socially selective processing.
An alternative explanation could be the difference in paradigms, as described in Chapter 3. While in the study by Jones and colleagues (2015), the social and non-

social condition involved the infants deliberately and actively choosing to visually attend to one or the other stimulus, while inhibiting the other, possibly recruiting frontal brain regions more than in the context of being presented with just one stimulus option.

The present study further adds to previous findings suggesting that at 6 months, differential theta activation can be measured to live social versus non-social stimulation, further strengthening the idea that by the age of 4-6 months, the infant brain has already undergone specialisation to social stimuli (Grossmann et al., 2008; Johnson, Grossmann & Kadosh, 2009; Johnson et al., 2005). Of note, theta power was not enhanced in a screen-based social context at that age (Jones et al., 2015), suggesting that by this age, theta power has specialised to social cues presented in a live context only, while by 12 months, the effect was also observed in the screen context. This might be explained by the fact that the live context offers to the child a multitude of multimodal social cues, which only in combination are strong enough to elicit differential activation in the specialising brain at this early age, in line with the theory of natural pedagogy (Csibra and Gergely, 2009).

This finding also shows that the effect is robust across paradigms, even at the younger age of 6 months. In Jones et al., 2015, the experimenter was singing and operating the toy at the same time, and the infant was free to look at either the face or the toy. Social and non-social condition were then defined by sequences of the infants' looking at either the experimenter or the toy. By contrast, in the present study, the social and non-social conditions were presented in separate, consecutive blocks of either signing *or* toy operating. The robustness of the effect across these paradigms points to its stability over *contexts* which also differ in their affordances to the child: In Jones et al., the child had to actively shift their attention to one or the other stimulus, while in the present study, one stimulus at a time was displayed. Further, the robustness of the effect across paradigms suggests that it is actually the *attending to* the stimulus that drives the effect, as opposed to just finding themselves surrounded by different contexts: In Jones et al., the experimenter was singing and looking at the child throughout, but theta power was only elevated in moments when the child was actually looking at the experimenter.

Finally, in line with the findings from the 14-month timepoint of this study, and consistent with previous findings from 6-month-olds (Jones et al., 2015), the amount of looking at the live stimulus did not differ between the social versus non-social condition, adding support to the idea that neural measures might be a more sensitive to capture differences between social and non-social live processing in infancy.

However, in contrast to the study by Jones and colleagues (2015), the effect was most prominent over posterior, and less strong in frontal regions. One reason for this difference in between studies might be the slight variation in paradigms. It is possible that infants recruited the frontal region more in the study in which they had to guide their attention between social and non-social input, while this was not necessary in the current study. Theta power over the frontal region has been shown to be increased during processes that require executive, top-down control of their attention (Anderson, Perone & Gartstein, 2022; Cavanagh & Frank, 2014; Orekhova, Stroganova, Posikera & Elam, 2006), which strengthens over the second half of the first year of age (Bazhenova et al., 2007; Orekhova et al., 2006). Also, as mentioned in Chapter 3, the current sample is more heterogeneous than the one reported in Jones et al., 2015, as it consists of infants with both typical and elevated likelihood for neurodevelopmental conditions; small-scale differences between groups might play a role in how different scalp regions are involved in social versus non-social processing. Thus, it might be that on a small-scale level the social-non-social difference is reduced in infants with elevated autism likelihood particularly over the frontal region, levelling out the whole-group social-non-social effect over this region. One recent study looking into the development of the theta network across the first year of age reported increasing connectivity during social versus non-social video processing throughout the brain network, instead of being limited to few regions (Van der Velde et al., 2021). The authors suggested that joint frontal and parietal theta activity might project into other parts of the brain which might explain the difference in region effects between studies (Van der Velde et al., 2021).

One possibility for future investigation would be to look at more fine-grained dissection of theta in relation to social versus non-social naturalistic experiences.

Different frequencies of the theta band might be involved in different functions. In 7- to 8.5-month-old infants, the lower frequency range of the theta band (4-5 Hz) was increased while anticipating the reappearance of a hiding experimenter in a live peek-a-boo game compared to baseline, as well as while looking at the experimenter after having reappeared compared to baseline. In addition, the higher frequency range of the theta band (5-6 Hz) were increased during the anticipation phase compared to both baseline and reappearance phase, possibly reflecting the additionally heightened attentional stage during anticipation (Stroganova et al., 1998). No interactions with region were observed. In a later study using a similar live paradigm (Orekhova, Stroganova & Posikera, 1999), the authors analysed power in the single 0.4-Hz-bins of the theta range (3.6-6.0 Hz) over different cortical regions. They observed increased power in several frequency bins over several regions during reappearance compared to baseline, and in even more frequency bins and over more regions during anticipation compared to baseline. In a slightly different live paradigm, 8-12 months old infants were a) looking at an experimenter speaking (social), b) actively exploring unfamiliar toys (toy), or c) looking at soap bubbles (baseline). Theta power was stronger in the 3.6-5.6 Hz range during social versus baseline and during toy versus baseline. While in the lower theta range (4.0 and 4.4 Hz) the effect of stronger power during social versus baseline was mainly observed frontally, in the higher frequency range (5.6 Hz), the effect was mainly observed over temporo-parieto-occipital regions (Orekhova, Stroganova, Posikera & Elam, 2006). It remains subject to further research whether specific parts of the theta band contribute differently to elevated theta power during naturalistic social experiences.

4.4.2 Differences by familial autism/ADHD status

Overall looking at the stimulus was reduced in infants at elevated likelihood for autism. This effect was only marginally significant in the present sample, and it was not observed at the 14-month-timepoint of the cohort (Chapter 3), and neither in the study by Jones and colleagues (2015). This finding was not expected and should be investigated in future studies, to see whether it is a true effect. It might be that characteristics of the paradigm play a role for this finding, since in both social and non-social condition the child had to look towards the experimenter in order to see the stimulus. Younger infants with emerging autism might direct their attention through the room differently, in that they look less towards the experimenter in general and hence to the toy located in the experimenter's hands. By contrast, other features in the room might capture their attention more than it is the case for typical infants. On the other hand, in this scenario the effect should have shown up in the study by Jones and colleagues (2015), too. It is hence more likely that the present marginal effect of reduced overall looking at the stimulus in infants with elevated autism likelihood is not a true effect which would probably not be replicated. The alteration in looking at the stimulus proportion in infants at elevated likelihood for autism was observed across stimulus conditions and was not specific to social or non-social stimulation. This is in line with what I observed a the 14-month-timepoint and what previous studies had reported (e.g., Jones et al., 2015).

Overall theta power was *reduced* in infants at elevated likelihood for ADHD, but only if they were at *typical* likelihood for autism. Reduced overall theta power in infants with elevated ADHD likelihood at 4-6 months is consistent with the finding at 14 months in the same cohort (Chapter 3), and with a recent study of the same cohort at 10 months of age (Begum-Ali et al., 2022). As theta power might signal cognitive control, reduced overall theta in the present cohort might reflect the onset of cognitive control difficulties that are characteristic for ADHD later in childhood. However, this pattern in the present cohort is contrary to what has been observed in younger infants and older children and adolescents. Infants of mothers with ADHD at six months of age (Shephard, Fatori, et al., 2019) as well as older children with ADHD (Barry, Clarke, Johnstone, McCarthy, & Selikowitz, 2009) exhibited *greater* overall theta power, counter to what one would expect. Further, elevated theta power may index a higher genetic likelihood for ADHD in adolescents (Tye et al., 2014). Interestingly, the effect in the present sample depended on concurrent autism likelihood: it was only present if there was typical likelihood for autism, suggesting the possibility that concurrent autism likelihood might be a protective factor for infants at elevated ADHD likelihood. More studies of independent samples are needed to further investigate this interaction.

The stimulus effect on theta power did not significantly differ by autism likelihood. This finding is contrary to my predictions. In the same infants, at 14 months, I did

observe an alteration by likelihood, with reduced theta power elevation during social versus non-social stimulation in infants with elevated autism likelihood (Chapter 3). Others observed this effect in toddlers with an autism diagnosis (Dawson, Bernier, et al., 2012)Because studies into neurodevelopment using other measures (fNIRS, ERPs) and paradigms (pre-recorded stimuli) showed alterations in infants with elevated likelihood at an age as early as 4-6 months (Jones et al., 2016; Jones et al., 2017; Lloyd-Fox et al., 2013), I had expected to see these alterations on the level of theta power, too. Therefore, it seems like alterations in theta power in infants with elevated autism likelihood come online only in the second half of the first year of age.
Crucially, the separate analyses of 14- (Chapter 3) and 5-months-old infants (in this chapter) do not allow to conclude that the reduction in the social/non-social effect on theta power in infants with increased familial likelihood for autism *significantly differs* between the 14- and 5-months-timepoint. However, they do show that the effect appears at 14 months, but not at 5 months. A next step would be running a secondary analysis of the initial models that showed a significant difference at one and not the other age point, including age as factor.

However, although the interaction between autism likelihood and stimulus condition was not significant, visual inspection of the data suggests that the stimulus modulation is reduced in infants with elevated autism likelihood. Especially when investigating regions and hemisphere-sides separately, it becomes visible that reduced elevation of theta power during social versus non-social stimulation in infants with elevated autism likelihood appears in virtually all regions and sides (although the stimulus X autism likelihood X region interaction was not significant). Interestingly, at 5 months, the modulation by likelihood seems weakest over posterior, where the stimulus effect was strongest, while primarily being visible over frontal, parietal and temporal regions. This reduced social-non-social difference in infants with elevated autism likelihood over the frontal region might also be responsible for the weak elevation of frontal theta power by social versus non-social stimulation compared to the strong difference over the posterior region.

Infants with elevated versus typical ADHD also did not differ significantly in theta power during social versus non-social processing. Visual inspection comparing likelihood groups across regions and hemisphere-sides (A3) suggests that at 5 months, the difference between theta power during social and non-social processing is also reduced in infants with elevated ADHD likelihood; while the reduced difference is still more prominent in the infants with elevated autism likelihood, especially over central electrodes.

Overall, the present results support earlier findings (Jones et al., 2015) that theta power during naturalistic experiences does distinguish between social and non-social attention at 4-6 months. They further indicate that the reduced effect in infants with elevated likelihood of autism observed at 14 months is not yet significant at 4-6 months, while visual inspection indicates already small differences that likely spread out and become stronger over the second half of the first year of age.

4.4.3 Differences by autism outcome

Infants with later autism diagnosis did not show statistically significant alterations in the social/non-social difference in theta power. Also, graphical examination did not suggest reduced theta power modulation in infants who go on to develop autism. This is contrary to what I had expected, since first, toddlers with an autism diagnosis did show these alterations, and second, infants with later diagnosis have shown alterations in social processing already around 6 months (Jones et al., 2016; Lloyd-Fox et al., 2018). However, this finding has to be interpreted with caution, since the sample of infants who went on to be diagnosed only consisted of 10 infants. It is likely that the variation in this small group was too large in order to reveal a potential effect. The same holds for the marginally significant effect of stronger overall theta power, and the finding of less overall looking at the stimulus in infants who went on to be diagnosed; they have to be read with caution due to the very small sample size. A larger group of later diagnosed infants would be needed to investigate the relation between live social/non-social theta power at 5 months and autism diagnosis at 36 months.

Interestingly, also continuous autism trait measures at 36 months were not related to differential theta power during naturalistic experiences at 5 months. I had expected to find a relation with continuous trait measures even if no relation with diagnosis can be established, but this was not the case. In fact, graphical inspection of the data did not reveal particular associations. It might be that at this age, differences in theta power during social versus non-social naturalistic experiences are not big enough yet to relate to later behavioural differences.

The specificity check with a continuous ADHD measures at 36 months revealed that indeed ADHD traits at 36 months were not related to the social versus non-social looking or theta power in infancy. However, I observed reduced *overall* looking at the live stimulus being related to higher ADHD scores at 36 months, which is in line with previous literature reporting higher overall looking times at 3 months, but a lack of growth in looking times between 3 and 24 months, compared to typical infants (Miller, Iosif, Young, Hill & Ozonoff, 2018), possibly reflecting early deviations in attention and cognitive control.

4.4.4 Predicting the trajectory of looking at a face

Only greater *overall* theta power at 5 months was related to fewer looks to a face at 10 months. Increased looking at the experimenter (but not toy) at 5 months was related to more looks to a face at 10 months and at 24 months, while theta power during social versus non-social attention did not predict the trajectory of looking at a face. The difference in theta power between social and non-social attention at this age likely was not strong enough to reveal associations.

4.4.5 Future directions

Theta power has been shown to play a crucial role in attention-guided processes such as during social versus non-social naturalistic experiences. It thus reflects a promising measure for the use on the individual level to reveal different levels of engagement with different social and non-social live stimuli. Theta power therefore can serve as a target response in the NBO study using a naturalistic paradigm.

In order to in the future further optimise a measure that is reliable and predictive for behavioural outcomes on the individual level, it might be worth combining theta power with other features of the theta band. For example, the temporal dynamics of theta activity during face processing might provide further inside into alterations in infants with emerging problems of processing social information, as 6- to 10-month-old infants with later autism diagnosis showed reduced theta trial coherence during face processing (Van Noordt et al., 2022). Further, it might be worth analysing individual trajectories of theta power during social versus non-social live action along with behavioural measures of social development across multiple timepoints in infancy, to take a further step towards predicting trajectories of early deviations in neural processing of social versus non-social information.

4.4.6 Conclusion

The present findings support and extend previous findings. They replicate studies showing that theta power is stronger during social versus non-social live attention at around 6 months. They further suggest that the effect is starting to be reduced in infants with elevated autism likelihood, especially over frontal electrodes, while this effect did not reach significance in the present study. Theta oscillations seem to be a suitable candidate for measuring attention engagement with social and non-social stimuli over the second half of the first year of age.

CHAPTER 5: PROOF OF PRINCIPLE – USING NBO WITH TO INFANT EEG TO STUDY INDIVIDUAL PREFERENCES FOR FAMILIAR AND UNFAMILIAR FACES

This study has been submitted for publication:

Throm, E., Gui, A., Haartsen, R., da Costa, P. F., Leech, R., Mason, L., & Jones, E. J. (submitted). Combining Real-Time Neuroimaging with Machine Learning to Study Attention to Familiar Faces during Infancy: A Proof of Principle Study.

5.1 Introduction

Attention allocation to social stimuli in the first months of age is crucial for the development of the social brain network (e.g., Johnson, 2011). The caregiver's face is a reliable and important social stimulus in the first year of age, which infants spend a large proportion of time looking at (Jayaraman et al., 2015). The infant Negative central ERP component which has been shown to be involved in attention engagement (e.g., Reynolds, Courage & Richards, 2010), showed differential responses to a caregiver's face at 6 months (e.g., De Haan & Nelson, 1997). However, this response towards seeing a familiar compared to an unfamiliar face has shown to change over the first years of age, while findings are inconsistent regarding the direction and timepoint of this change. The present study uses the Nc/mother-stranger paradigm to apply Neuroadaptive Bayesian Optimisation (NBO) to infant EEG data for investigating individual differences in social attention engagement.

5.1.1 The Negative central (Nc) ERP component as neural correlate of infant attention engagement

The *Negative central* (*Nc*, Courchesne, Ganz, & Norcia, 1981) component of the infant ERP measured with EEG is a negative deflection most prominently occurring over frontocentral electrodes between 350 and 800 ms after stimulus onset. In various studies, it has been shown to be involved in attention engagement. Studies simultaneously recording EEG and physiological measures reported a higher Nc amplitude during physiologically (heart-rate) defined periods of attention versus inattention in infants, regardless of stimulus modality and type (Guy, Zieber & Richards, 2016; Richards, 2003; Richards, Reynolds & Courage, 2010). Further, the effect of greater Nc amplitude during attention versus inattention increases in strength between 4.5 and 7.5 months (Richards, 2003). One study investigating Nc amplitude responses across multiple time-points in infancy reported an increase in negative amplitude in the first year and a decrease in the second year (Luyster et al., 2014). Overall, these findings suggest that a stronger Nc is associated with deeper attentional engagement with a stimulus, but that this response changes across infancy.

Because the infant Nc component seems to be an index of attention, researchers have used it to study the stimuli infants are paying more attention to. It has been used in a wide range of paradigms using visual stimuli, primarily faces, and has been shown to be modulated by a number of factors. Often, the effect changes with age.

Faces versus objects. Multiple studies investigated the ability of the infant Nc amplitude and latency to differentiate faces from objects across different age points in infancy and early childhood, revealing rather inconsistent findings. By 3- to 4-years, children show a more negative Nc to objects than faces (Dawson et al., 2002). The differential Nc response to faces versus objects seems to develop only in late infancy. In young infants aged around 4-6 months, evidence is mixed with some studies reporting larger Nc amplitudes towards objects versus faces (Conte, Richards, Guy, Xie, & Roberts, 2020), while others reported no amplitude differences between faces and objects in a typically developing sample (Jones et al., 2016), stronger amplitude for objects versus faces in infants with later autism only (Jones et al., 2016), or trends towards stronger amplitude to faces versus objects at 4 months and objects versus faces at 6 months (Webb, Long & Nelson, 2005). In mid infancy, findings continue to draw a mixed picture. Some studies found at 6 and 8 months the Nc amplitude to be larger during attention to faces versus objects, and at 12 months, the Nc amplitude did not differ by stimulus and/or attentional state (Conte et al., 2020), while others did not observe an amplitude difference at 8 months (Webb, Long & Nelson, 2005). Towards the end of the first year of age, findings become more consistent reporting that at 8, 10 and at 12 months, the Nc amplitude did not differ by stimulus and/or attentional state (Conte et al., 2020; Webb et al., 2005). Regarding latency, Nc responses did not differentiate faces from objects at 4, 6, 8 and 10 months in typically developing infants (Jones et al., 2016; Webb et al., 2005), but were observed to be shorter in response to faces versus objects at 12 months (Webb, Long & Nelson, 2005), and had a later offset for objects versus faces in 6-month-old infants with later autism diagnosis (Jones et al., 2016).

Emotional facial expressions. Research on the Nc in response to emotional facial expressions in the second half of the first year has been mixed, too. Stronger responses were reported to fearful faces (e.g., Grossmann et al., 2011; Nelson & de Haan, 1996) or angry faces (e.g., Grossmann et al., 2007), suggesting stronger attention engagement with stimuli that bear a potential thread towards the individual and that the infant might not be often exposed to (but see e.g., Stahl et al., 2010, for contrary findings).

Saliency or Familiarity. Moreover, the Nc amplitude has been shown to be modulated by saliency or novelty (e.g., Dawson et al., 2002; Richards, 2003; Key & Stone, 2012; Luyster et al., 2011; Guy et al., 2016; Guy et al., 2018; Carver et al., 2003; Reynolds and Richards, 2005; De Haan and Nelson, 1997, 1999; Guy et al., 2013; Reynolds et al., 2010; Webb et al., 2005), with some studies reporting stronger Nc response to novel versus familiar stimuli (Reynolds & Richards, 2005), and others reporting stronger amplitude responses to familiar versus unfamiliar stimuli (e.g., De Haan & Nelson, 1997; Luyster et al., 2014; Webb et al., 2011). This inconsistency might be related to the fact that at different periods in development either the familiar or the unfamiliar stimuli appear more relevant and hence salient to the infant.

5.1.2 Nc responses to faces of parent versus stranger

The direction of the Nc amplitude difference between familiar and unfamiliar faces, but not objects, seems to change with age, supporting the idea that the Nc is modulated by the *relative* salience of the stimulus. At 6 months of age, infants showed greater Nc responses to their mother's face than a stranger's face (De Haan and Nelson, 1997, 1999; Webb et al., 2005), provided that the faces were looking dissimilar (De Haan & Nelson, 1997). However, this pattern seems to change with age in the first year of age, as well as with their socio-cognitive development and the unfolding relation with their caregivers.

5.1.2.1 Age-related changes in the Nc to parent versus stranger

The greater Nc amplitude to mother's face versus a stranger's face has been consistently reported for 6-month-old infants (De Haan and Nelson, 1997, 1999; Webb et al., 2005; but see Swingler, Sweet & Carver, 2007 for the reversed pattern, and Luyster et al., 2014 for no difference). Numerous studies have indicated a shift from greater Nc to mother's versus stranger's face towards greater Nc to stranger's versus mother's face towards the end of the first year and in early toddlerhood. By 3-5 years of age, children show greater Nc negativity towards stranger's compared to mother's face (Carver et al., 2003; Dawson et al., 2002; Moulson, Westerlund, Fox, Zeanah, & Nelson, 2009). However, findings as to when this shift happens are highly mixed.

Some studies observed a shift towards greater attention engagement with stranger's face not before the second year of age. At 12-17-months, typically developing children showed stronger Nc responses to familiar versus unfamiliar faces (Webb et al., 2011). Infants at 6-36 months of age showed a greater Nc negativity for mother's versus stranger's face (Luyster et al., 2014). Until 24 months, the Nc was greater for mother's face (Carver et al., 2003). Other studies reported stronger Nc amplitude for stranger's versus mother's face already in infancy, at 9 months (Key & Stone, 2012) and at 12 months (Guy, Richards, Tonnsen & Roberts, 2018; Luyster et al., 2011). A recent study did not observe Nc amplitude differences towards parent versus stranger in 12-month-old infants, possibly because individual differences in whether or not the shift has already occurred overlaid group-level findings (Glauser et al., 2022).

The shift in the direction of the Nc amplitude difference towards familiar versus unfamiliar faces might reflect the degree to which these faces are salient in the respective period of development in early childhood, with a focus on relationship with the primary caregivers in infancy and early toddlerhood, shifting towards more exploring the greater social surrounding including new faces later on. Besides age, the pattern of Nc responses to familiar versus unfamiliar faces also seems to depend on social development of the infant.

5.1.2.2 Modulation of the Nc response to parent versus stranger by social development

Typically developing 2-5-year-old children showed stronger Nc amplitude towards stranger versus parent, while in children with autism, the Nc did not differentiate between familiar and unfamiliar faces, but only between familiar and unfamiliar objects (Dawson et al., 2002), suggesting the Nc to reflect impaired face recognition in children with social behaviour difficulties. Further, a study pooling together data from 6-to 36-month-old infants observed a stronger Nc amplitude response for mother's versus stranger's face in infants at typical but not elevated familial likelihood for autism (Luyster et al., 2014). Further, while 12- to 17-month-old typically developing children and 18- to 30-month-old children with autism showed stronger Nc responses (mean amplitude between 350-750 ms) to familiar versus unfamiliar faces, this was not observed in the 18- to 30-month-old typically developing children

(Webb et al., 2011), suggesting delayed development in infants with autism, with differences being observable already in the second year of age.

Going further back to the end of the first year, infants with elevated likelihood for autism might differ in when they shift to showing more attention engagement with the unfamiliar versus familiar face. At 9 months (Key & Stone, 2012) and at 12 months (Luyster et al., 2011), infants with increased familial likelihood for autism showed stronger Nc amplitudes towards the stranger's versus parent's face, just as did the typical infants in these studies. Of note, the proportion of 12-month-olds showing the greater Nc response to stranger's versus mother's face was larger in the typical likelihood group than in the elevated likelihood group (Luyster, Wagner, Vogel-Farley, Tager-Flusberg, Nelson, et al., 2011), suggesting that more infants in the typical versus elevated likelihood group have already shifted towards stronger attention to unfamiliar faces. Consistent with this idea, in one study, 12-month-old infants with elevated likelihood for autism did not show the stronger Nc amplitude towards stranger's versus mother's face observed in the typical infants, and this was also the case for unfamiliar versus familiar objects (Guy et al., 2018). Of note, this difference was not observed on the level of heart rate, highlighting the importance of neural measures in this context (Guy, Richards, Tonnsen & Roberts, 2018). Finally, a recent study with 12-month-old infants reported that neither infants with nor without later autism diagnosis showed differential Nc amplitudes towards parents versus stranger's face, potentially indicating individual differences overlaying group effects (Glauser et al., 2022). Overall, these findings suggest that the stronger Nc amplitude to familiar faces in infancy, the timepoint of an attentional shift towards unfamiliar faces, and a stronger Nc amplitude to unfamiliar faces later in infancy and toddlerhood might differ in children with elevated familial likelihood for autism.

In line with altered responses in children with autism diagnosis or elevated likelihood, the Nc amplitude to a stranger's versus parent's face was also observed to differ by continuous measures of social development. In 6-month-olds, stronger Nc amplitudes to stranger's versus parent's face were related to more proximity- and interaction-seeking behaviours during separation and reunion with a parent, suggesting that infants who become more initiative and active in interaction with the caregiver show the shifted pattern of stronger amplitude towards strangers versus parents face observed in older children (Swingler, Sweet & Carver, 2007). These results suggest that the developmental shift in Nc responses towards stronger amplitudes for stranger's versus parent's face is related to how the child interacts with the parent and to their unfolding relationship. At the same time, stronger Nc amplitudes towards *mother's* versus stranger's face were related to increased infant distress in 6-month-olds (Swingler & Carver, 2013), but not ot related to interpersonal skills measured by the VABS-II, while shorter Nc latencies to the parent's but not stranger's face were related to stronger interpersonal skills measured by the VABS (Key & Stone, 2012). Finally, Nc amplitude to mother versus stranger also seems to depend on the child's level of cognitive development. At 9 months, the Nc amplitude to mother versus stranger was larger in infants who were quicker learners earlier in infancy (Reeb-Sutherland, Levitt & Fox, 2012), and at 12 months, greater Nc amplitude towards mother versus stranger in typically developing 12-month-old infants was related to higher scores on the expressive language subscale of the MSEL (Glauser et al., 2022).

5.1.2.3 Summary

A more negative Nc amplitude has been indicated to reflect greater attention engagement and be modulated by saliency of a stimulus. Around the sixth month, findings consistently show larger Nc amplitude for familiar versus nonfamiliar faces. This pattern starts to change further in development. Findings are mixed as to when this change occurs, suggesting that there might be individual differences overlaying group-level Nc responses. In order to reliably measure attention engagement to familiar and unfamiliar faces in the individual infant, the present study used NBO to identify the stimulus maximally eliciting an infant's Nc response.

5.1.3 The present study

5.1.3.1 NBO to map individual preferences to parent versus stranger

The present study is a proof of principle study combining NBO with infant ERP data to study individual infants' attention engagement, particularly the function of the Nc response to images of faces of the own parent versus a stranger.

EEG was recorded while infants aged 5 to 12 months were presented with images of faces resulting from morphing parent's face into stranger's face. The faces were ordered along a continuous stimulus space for the Bayesian Optimisation to sample across, with the parent's face and a stranger's face at its extremities. The Nc mean amplitude was calculated in real time after repeated presentation of one face and passed to the Bayesian Optimisation algorithm aiming to identify the stimulus that reliably produces the strongest Nc response in an individual infant. The optimal stimulus was defined as identified when the Bayesian Optimization converged, that is kept sampling this same stimulus. Identifying the most engaging stimulus in the individual allowed us to link individual differences in social attention engagement to age, behavioural and environmental characteristics.

By mapping an individual's response function to a stimulus space, NBO can be a useful tool to bridge the gap between group-level and individual-level reliability of infant brain responses to social stimuli, such as the Nc response to parent-stranger-faces. Instead of repeatedly presenting the face of either parent or stranger, the current design allowed to present faces linearly changing from parent's face into stranger's face by sampling a 1-dimensional continuous stimulus space of morphed faces of the parent's and stranger's face, with parent's and stranger's face at its extremes.

Infants were tested at the age of 5-12 months, targeting the age range across which according to group studies developmental change developmental change occurs in the Nc response to parent versus stranger. The relatively wide age range allowed to make age group comparisons of the individual optima obtained in the different sessions, and to see whether the number of infants "choosing" one over the other optimum differs between younger and older infants.

Previous studies have demonstrated a significant difference in Nc between mother's and stranger's faces was only found for infants presented with a stranger's face that looked dissimilar from their mother's face (De Haan and Nelson, 1997). In order to account for a possible effect of similarity in this study, two independent researchers judged the similarity of each parent-stranger pair of faces.

5.1.3.2 Hypotheses

To serve the **proof-of-principle goal** of the study, I hypothesised that the method works practically, that it yields reliable results, and that it yields valid results.

To evaluate whether the method works *practically*, I predicted that attrition rate would be lower than in traditional infant EEG studies where attrition is 21-23% on average. In classic infant ERP studies, attrition unrelated to experimental error reaches 23% in 5-month-olds and 21.3% in 10-month-olds (van der Velde & Junge, 2020). Due to the greater variety in stimuli as well as the presentation being guided by the individual's attention, I expected lower attrition in the present experiment.

To evaluate whether the method yields *reliable* results, I predicted that the NBO would reach the early stopping criterion before reaching 15 blocks in the majority (>50%) of the infants completing the paradigm, indicating that responses mapped on the search space are reliable, such that stronger signals were consistently concentrated in the same region of the stimulus space of one individual (Da Costa et al., 2021). The goal of 50% of infants completing the paradigm reflects the half of the infants aged 5-8 months for which a preference for parent versus stranger was expected. For infants aged 9-12 months, I did not have a clear prediction as no clear preference towards one point of the space can result in non-convergence just as unreliable responses can. Of note, reliability of the empirical values is embedded by design, as the early stopping criterion can only be reached if there is low variability (i.e. high reliability) of the empirical values sampled at a given point in the space.

To evaluate whether the method yields *valid* results, that is that optima really represent what they are supposed to measure, it was predicted that the Nc amplitude in infants with optima closer towards the parent were indeed stronger for parent versus stranger, and in infants with optima closer towards the stranger stronger for stranger versus parent.

On the contrary, a clear sign of failure of the application of the method would be if the algorithm did not converge in most infants. Reason for this could thereby either be that the sampled target values were not reliable, preventing the uncertainty in the predicted function to decrease and prevent the algorithm from exploiting a space region where maxima are predicted (both because values are not consistently high, and because high uncertainty across space makes algorithm continue exploring different regions of the space). Further, if a subgroup's Nc amplitude direction does not match what the individual optimum indicate (e.g., Nc parent>stranger in infants converging toward the parent side), this could raise doubts with regard to the validity of the BO optima, that is whether they really represent points at which the greatest Nc is expected, or are just arbitrary points in the space instead.

Of note, comparing the results with previous results to evaluate the success of the novel method applied to infant EEG is not useful with this mother-stranger paradigm, as previous group-level findings were mixed and there is no ground truth (which holds for many of the infant literature; instead, the individual-level results can help explain the pattern of inconsistent and null-effect group-level findings). The present NBO study though has the potential to disentangle the reasons for this inconsistency in findings by studying results on the individual and subgroup level.

To serve the **theoretical goal** of the study and use the advantages of the novel approach to gain new insights on infant social attention, I linked individual optima to age, behavioural and environmental characteristics.

I hypothesised, first, that the position of the optima in the search space were related to age and social behaviour. I expected optima to be closer to the familiar versus unfamiliar face in the stimulus space in the younger infants (5-8m), and no preference or preference for stranger's face in the older infants (9-12m), based on reports of stronger Nc towards the familiar face in infants aged around 6 months (De Haan & Nelson, 1997; De Haan & Nelson, 1999; Nelson et al., 2000). Based on previous research showing greater Nc responses to familiar versus unfamiliar faces in autistic children aged 18-30 months (Webb et al., 2011), resembling the pattern observed in typically developing children at an earlier age, I predicted that optima closer to the parent's face are related to lower scores on the VABS Socialisation Domain. I further predicted a modulation of the optimum stimulus by indices measuring the infant's parent-reported interest in and distress towards familiar and unfamiliar people, with an optimum towards the mother's face to be related to higher interest in familiar people as well as lower interest in and higher distress towards new people.

Finally, to compare individual-level with group-level results, I calculated the Nc amplitude in the traditional, group-based way in response to "pure" parent vs "pure" stranger, ignoring the mixed images created by morphing the two originals. In line with the individual-level predictions, I hypothesised the Nc amplitude to be larger for the familiar compared to the unfamiliar face in the younger but not older infants, and in infants with higher socialisation scores, higher interest in familiar persons, lower interest in new persons and lower distress towards new persons. To control for a possible effect of caregiver-stranger-similarity, analyses were re-run including similarity as a covariate.

5.2 Methods

5.2.1 Participants

In total, 62 infants (27 females, 35 males) aged 5 to 12 months (m=266.69 days, SD=61.88 days, range=120-375 days) took part in the real-time EEG experiment. The younger age group (n=30, 13 females, 17 males) was aged 5 to 8 months (m=214.07 days, SD=39.03 days, range=120-265 days), and the older age group (n=32, 14 females, 18 males) was aged 9-12 months (m=316.03 days, SD=30.25 days, range=272-375 days). Children were not invited if they had a family or personal history of epilepsy, if they were born pre-term (<31 weeks of gestational age), if they had a sensory or motor impairment or any clinical condition.

5.2.2 Real-time EEG task

5.2.2.1 Stimuli

Stimuli were the face of the infant's accompanying caregiver and a gender-matched stranger, presented on a screen. Faces were centred on the image and the body below the neck was cropped; the emotional expression was neutral. The same images of male and female stranger's face, respectively, were used across all infants. Before each testing session, the software StyleGAN2, a deep learning algorithm for generative image modelling (Karras, Laine, & Aila, 2019), was used to morph the parent's and stranger's image into each other in order to produce 10 additional, realistic images showing the respective parent's face linearly changing into stranger's face (Figure 5.1). The total 12 stimuli were arranged in a stimulus space, varying in the dimension of similarity to parent's face, with the parent's face and the stranger's face as extremes of the continuum. Based on previous research (e.g., de Haan & Nelson, 1997, 1999), I hypothesised that the target brain response (mean Nc negativity) would be maximised by mother's face in the younger age group.

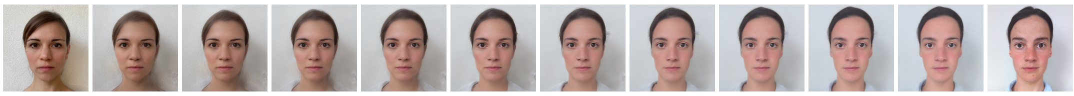


Figure 5.1. An example of the 1D parent-stranger stimulus space

To account for a possible effect of similarity between the stranger's face and the respective parent's face reported previously (De Haan & Nelson, 1997), two independent researchers blind to the study design judged each pair of mother-stranger faces based on facial configuration and qualitative aspects like hair colour. Similarity was rated by two independent raters on a continuous rating scale using a response slider, ranging from 0 (very dissimilar) to 100 (very similar), and inter-rater agreement was calculated using Pearson correlation as in De Haan & Nelson (1997).

5.2.2.2 Procedure

The infant sat on the parent's lap approximately 60 cm from a 24-inch diagonal screen. The entire paradigm consisted of a maximum of 15 stimulus-presentation blocks, with each consisting of 12 trials of the same face, aiming for least 10 good trials per block to produce the ERP and considering a data loss of 15% due to bad data quality (e.g., Munsters et al., 2019).

Stimulus presentation was implemented in MATLAB using PsychToolBox-3. Each block started with the presentation of a spinning red spiral to attract the infant's attention to the screen. The experimenter was monitoring the infant's gaze via a webcam located on top of the participant screen. Once the infant was looking at the screen, the experimenter elicited the start of the first trial per key press. After a fixation cross (500-1000 ms), the face image was presented on a grey background for 500 ms. Then, the next trial (fixation cross + image) started, given the infant was still looking at the screen. Stimulus presentation during the block continued until the infant was looking away from the screen. If the infant was looking away, the red spiral, paired with a brief sound, appeared on the screen, until the infant was looking back at the screen. If the infant's gaze was still diverted from the screen, a set procedure was followed aiming to bring the infant's attention back to the screen. An experimenter standing out of sight during the beginning of the paradigm would 1) give a teething ring to the baby; 2) sit next to the infant, tap on the screen, saying "<Baby's name>, look!"; 3) offer a soft snack. If the infant would still not look back at the screen, the experimenter would 4) ask the parent for advice on what could help their baby focus on the screen and implement respectively. As a last resort, the experimenter would 5) blow bubbles in front of the screen. After every few trials, for a total of three times, a brief (~2 seconds) colourful cartoon video paired with infant-friendly music was presented as dynamical attention grabber in order to reward the infant for looking at the screen and to further enhance their attention to the screen, as well as to reduce habituation to the faces. After the 12 face trials of a block, a cartoon still image was presented on the screen, while the EEG data obtained during the block was processed in real time and the output value passed on to the BO to select the next face stimulus to be presented. During this period (~ 6 seconds), the infant did not have to look at the screen, but could move and look around, while the experimenter only minimally interacted with the child so that the following stimuli were not perceived boring in comparison.

5.2.2.3 EEG data acquisition (as in Throm et al., 2023)

Lab Streaming Layer (LSL) was used to stream the EEG data and read it into MATLAB during the experiment. EEG data was recorded using the gel-based, wireless ENOBIO 8-channels EEG system (NE Neuroelectrics; 10-10 EEG coordinate system) with 6 fronto-central electrodes of interest (Fz, FC1, FC2, C1, C2 and Cz) and two reference electrodes (P7 and P8; Figure 5.2). I used skin-adhesive stickers to attach behind the ear the system's two default electrodes for online referencing of the signal (common mode sense, CMS; driven right leg, DRL). EEG data was digitized at 500 Hz. Before the start of the experiment, signal quality was assessed visually in the acquisition software, including impedance, noise and drift of each of the eight electrodes.

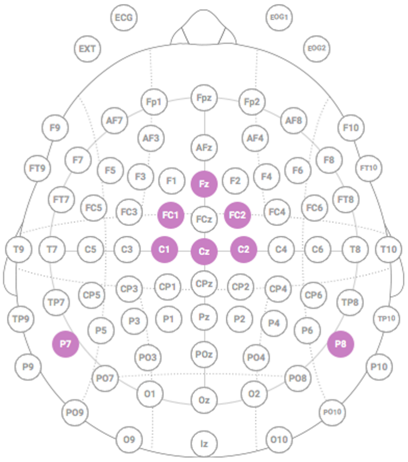


Figure 5.2. Electrode mount including six channels of interest (Fz, FC1, FC2, Cz, C1, C2) and two reference channels (P7, P8)

5.2.2.4 EEG data pre-processing

Segmentation, filtering, baseline correction. After each stimulus presentation block (=12 repeated presentations of 1 face), the streamed EEG signal was read into MATLAB, pre-processed and the mean amplitude of the negative deflection of the Nc component calculated using custom MATLAB scripts. The raw EEG data was cut in 1500-ms-segments around the stimulus marker. Segments were detrended, demeaned, mirror-padded (padding value: 1000) and band-pass filtered (0.1 to 20 Hz). Smaller segments representing the time-window of interest around stimulus onset were extracted from the segments (100 ms before to 800 ms after stimulus onset). The data post stimulus onset was subtracted by the data recorded pre-stimulus onset, in order to correct for the baseline signal.

Artifact rejection. Each trial in each channel was checked for artifacts. If an artifact was detected, this trial for this channel was excluded from the further analysis. An artifact was identified if within the time-window of interest a) the signal exceeded a threshold in amplitude or b) a threshold in range or c) if the signal was consistently flat (below 0.0001 µV). In order to account for individual differences in the magnitude of the ERP signal between infants, the amplitude and range thresholds were determined for each individual at the beginning of the experiment. To do this, the script calculated the Nc obtained after the first block with more lenient thresholds (amplitude: +/-250 µV; range: of +/-500 µV). If the peak amplitude of this Nc was below +/-*200* µV in amplitude, the script then used more conservative thresholds for artifact rejection in this particular baby (amplitude: +/-200 µV; range: of +/-400 µV). In that case, the scripts re-calculated the ERP from the first block using the adapted thresholds, before passing on the output value to the BO. The individualised thresholds ensured that for infants with naturally larger ERPs no data is excluded, while for infants with smaller ERPs all artifacts can be excluded.

Re-referencing and averaging. For each clean trial of the channels of interest (Fz, Cz, FC1, C1, FC2, C2), the signal was subtracted by the pooled signal of the reference channels P7 and P8. After re-referencing, the signal was averaged across all samples in all channels of interest.

Re-referencing was applied because re-referencing revealed to lead to a stronger difference in Nc mean amplitude between parent and stranger when comparing the responses to mother versus stranger in eight pilot infants (here: re-referenced to P7). This was likely due to re-referencing allowing to cancel out common noise affecting all electrodes, in turn increasing the signal-to-noise ratio (Luck, 2014). P7 and P8 were chosen as re-reference electrodes, because the alternatives were either not feasible with the system and electrode arrangement used, or less favourable in terms of data survival after artifact rejection. As such, the overall average of electrodes was not used for rereferencing, in order to not cancel out signal of interest, given that all 6 of 8 electrodes available were used as electrodes of interest. The linked mastoids were not used for re-referencing (Luck, 2014) because using the ENOBIO system required using one of the mastoids for attaching the online reference electrode. Using pilot data, I compared different alternative re-reference options including occipital and parietal electrodes. A total of 111 ERPs collected from 13 pilot infants were re-referenced to P7, Oz and the average of P7 and Oz, and the amount of data that survived artifact rejection was compared between the re-referencing methods. When re-referencing to Oz or the average including Oz, only 50% of the trials across channels of interest survived the artifact rejection step and could hence be included in the ERP. By contrast, when re-referencing to P7 only, 60% of the trials survived artifact rejection. With the aim of preserving as much data as possible, the more lateralized P7 therefore seemed to be a better re-referencing candidate than Oz, presumably because it was less affected by the infants' leaning back at their parents' chest and hence less prone to noise. To avoid lateralization effects when using the left lateralised P7, average of P7 and P8 was used as electrodes for rereferencing.

5.2.2.5 EEG metric for the Bayesian Optimisation

After pre-processing, the EEG output value was calculated to be passed on to the Bayesian Optimisation. In the present study, the Nc mean amplitude was used because it was shown to reflect attention engagement (Conte, Richards, Guy, Xie & Roberts, 2020b; Gui et al., 2021; Richards, 2003)). The Nc mean amplitude was defined as the mean amplitude of the biggest negative deflection within the typical Nc time window of 250 ms[2] to 800 ms.

Calculation steps to obtain the Nc mean negativity were scripted in MATLAB, too. First, the point of the lowest amplitude of the signal within the Nc time window was identified as the peak of the Nc. If this peak was positive, that is no deflection below zero was present, the mean amplitude across the entire Nc time window, instead of only of the negative deflection, was used as an alternative measure in the respective block. If this peak was negative, the onset (= first crossing of x-axis before peak) and offset (= first crossing of x-axis after peak) of the negative deflection around this peak was determined. If the onset or offset point was located beyond the boundaries of the Nc time window (250-800 ms), the respective boundary of the Nc time window was used as onset or offset.

Previous literature indicated that in order to detect significant differences between face and non-face stimuli within the Nc time window, the duration of the negative deflection of the ERP has to pass a minimum threshold of 58 ms (Gui, 2019). Hence, as a sanity check, the duration between onset and offset was calculated, and if it did not meet the threshold of 58 ms, the mean amplitude across the entire Nc time window, instead of only of the negative deflection, was used as an alternative measure in the respective block. Otherwise, if the duration between onset and offset passed the threshold of 58 ms, the mean amplitude of the signal within the defined onset or offset point was calculated and used as EEG output value to be passed on to the BO.

5.2.2.6 Online EEG data quality check

Before passing on the EEG output value to the BO, real-time EEG data quality checks were performed in MATLAB. First, the percentage of valid trials from all recorded trials across the channels of interest (6 x 12 trials = 72 trials) was calculated. In order for the ERP output value to be passed to the BO, it had to include at least 10 (=~ 15%)

---

[2] The typical Nc time window is 300-800 ms post stimulus-onset. In the present study, I adjusted this time window to take into account the lag of the streamed EEG data relative to the marker time stamp. This lag was in the pilot phase of the study estimated to be ~35 ms.

valid trials. If this threshold was not met, the ERP output value was not passed on to the BO, and instead the entire block was repeated.

As additional real-time data quality check, the number of valid trials per channels was plotted after each block (Figure 5.3). This allowed the researcher to identify potential channels of poor quality, giving the opportunity to undertake adjustments on the cap to improve the signal quality.

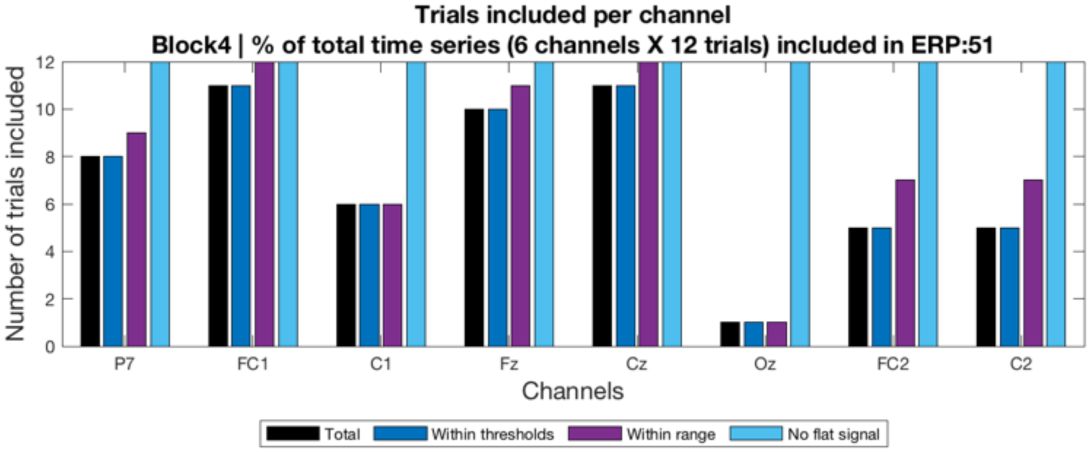


Figure 5.3. Bar chart displaying the number of valid trials per channel in the current block (Da Costa et al., 2021)

5.2.2.7 Eye blinks

Besides making sure that the infant was looking at the screen during stimulus presentation, eye blinks were not controlled for in real-time, because their influence was expected to be negligible (as mentioned in section 2.8.4.3 Dealing with eye blinks).

5.2.2.8 Bayesian Optimisation

In the current study, the Bayesian Optimisation algorithm was programmed towards maximising the Nc negativity by sampling towards the optimum point in the parent-stranger stimulus space. Of note, because of the negative valence of the Nc deflection, I aimed to elicit the most *negative* Nc mean amplitude value.

I used 4 burn-ins to the experiment. Burn-in points were the extremes of the 1-dimensional search space as well as two images from between, so that the spatial distance between each burn-in image was constant (stranger's face – 1/3 parent's face – 2/3 parent's face – parent's face). The four burn-in images were presented to all infants in randomly assigned order.

As exploration/exploitation hyperparameter ξ, a value of 0.1 was used. This value had in pilot sessions revealed to allow in most infants for a fair amount of exploration before then rapidly converging to the predicted optimum.

Convergence was reached when the same point was sampled three consecutive times, indicating identification of the optimum (as in Lorenz et al., 2018),that isthe point predicted to elicit the maximum Nc negativity. If convergence was not reached, the paradigm would stop after 15 blocks. This number equivalents to approximately 20 minutes, which was the maximum amount of time I expected the infant to be attentive. For a more detailed description of the Bayesian Optimisation mechanisms and chosen parameters see Chapter 2. An example of the individual optimisation process of the current study is shown in *A.5.1*.

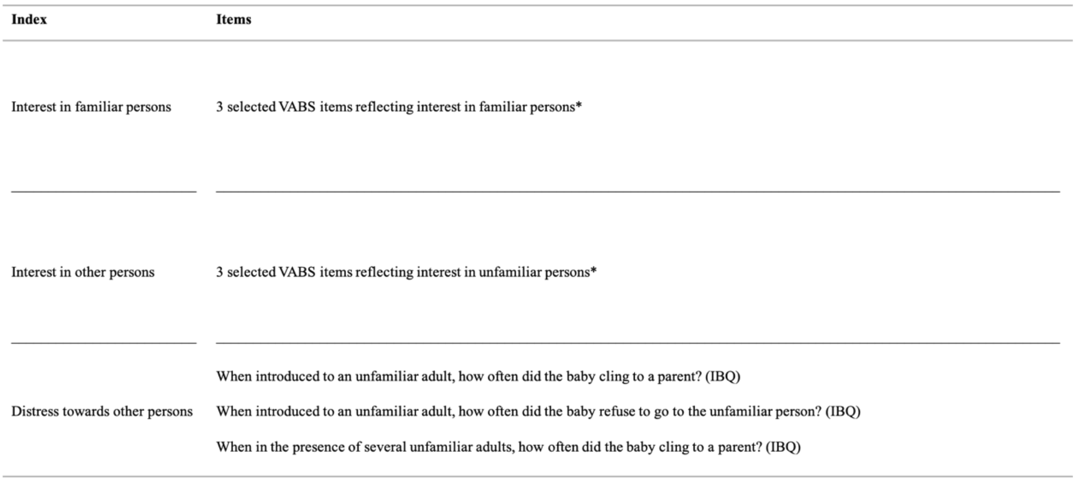
5.2.3 Parent-report measures of social development

Parents were asked to fill in online questionnaires before their visit to the lab. Questionnaires used in this study were the Vineland Adaptive Behavior Scale (VABS) and selected derived questionnaire indices derived from selected items from the VABS and the Infant Behavior Questionnaire-Revised (IBQ-R; Garstein & Rothbart, 2003). All questionnaires were administered online using the Gorilla Experiment Builder (www.gorilla.sc; Anwyl-Irvine, Massonnié, Flitton, Kirkham & Evershed, 2020).

Vineland Adaptive Behavior Scale (VABS). As described in Chapter 3, the VABS is a semi-structured interview measuring adaptive functioning in everyday life (Sparrow et al., 2005). The key behavioural variable was the VABS (Parent/Caregiver Form) Socialization Standard score, including the subdomains *Interpersonal Relationships* (10 items) and *Play and Leisure Time* (8 items). Items of a subdomain are administered from the age-specific starting point, and until the ceiling level is reached. Of note, as a *mistake in the present study,* for the Interpersonal Relationships subdomain, only items designed for infants below 1 year of age were administered. Therefore, the ceiling level was often not established. When the ceiling level was not established, I used the last item administered of the respective subdomain as ceiling level. Of note, this means that possible effects are more likely to not be detected, given the reduced variability; if on the other hand significant effects emerge, they can be interpreted. Further, in addition to the overall Socialisation Standard score, models were also run with the two subdomain scores separately, to be able to disentangle potential effects,

Interest/Distress towards familiar and unfamiliar persons. Besides overall socialisation skills, in order to evaluate the relationship between optimum-mother distance and behaviour towards familiar and non-familiar persons in specific, I extracted selected items from the VABS and the IBQ in order to create three indices measuring interest and distress towards familiar and unfamiliar persons: “Interest in familiar persons”, “Interest in new persons”, and “Distress toward new persons”. A list of the items composing each index can be found in *Table 5.1*. For each infant, the mean across the raw scores to each index’ items was calculated. These three mean values per child were used as a variable in the present study.

Table 5.1. Indices derived from items of the Vineland Adaptive Behavior Scales (Sparrow, Cicchetti & Balla, 2005) and Infant Behavior Questionnaire (Rothbart, 1981; Gartstein & Rothbart, 2003; Putnam et al., 2014) – revised to measure interest and distress towards familiar and unfamiliar people. *Of note, VABS items cannot be displayed due to copyright restrictions.



Index	Items
Interest in familiar persons	3 selected VABS items reflecting interest in familiar persons*
Interest in other persons	3 selected VABS items reflecting interest in unfamiliar persons*
Distress towards other persons	When introduced to an unfamiliar adult, how often did the baby cling to a parent? (IBQ) When introduced to an unfamiliar adult, how often did the baby refuse to go to the unfamiliar person? (IBQ) When in the presence of several unfamiliar adults, how often did the baby cling to a parent? (IBQ)

5.2.4 Statistical (offline) analysis

Overall attrition and convergence. As a measure of attrition, the proportion of infants who completed the paradigm was calculated. Further, efficiency of the algorithm was calculated by the proportion of infants for whom the algorithm converged, with convergence being defined as the reaching the early stopping criterion of the algorithm's sampling of the same stimulus for three consecutive iterations (for more information on stopping criteria, see Chapter 2).

Convergence towards parent versus stranger face. The primary outcome of interest after an individual's session was the position of that individual's optimum in the parent-stranger stimulus space. In additional statistical analyses after data collection of the entire sample was completed, the individual optima in the search space were analysed and the relation between the position of the individual optima and other measures (age, behaviour) was tested. The position of the optimum in the space was measured, first, as Euclidean distance from the parental face ("optimum-parent distance", continuous). A shorter optimum-parent distance reflected closer proximity of the predicted optimal stimulus to the parent's face in the parent-stranger-stimulus-space (i.e., more similar to parent). Second, the position of the optimum in the space was measured dichotomously as being closer either to the parent's or the stranger's face in the parent-stranger stimulus-space.

Relation with age. Overall convergence (yes/no) was analysed in relation to age in days using logistic regression. ANOVA was used to test whether the optimum-parent distance differed by age group (5-8 m versus 9-12 m), and linear regression to test whether the optimum-parent distance was predicted by age in days. Logistic regression was used to test the relation between convergence to parent versus stranger and age in days. Fisher's test was used to calculate the odds for converging for parent versus stranger by age group (5-8m versus 9-12m).

Relation with behaviour. Multiple linear regression and multiple logistic regression were used to test whether the optimum-mother Euclidean distance or the odds of converging towards parent versus stranger, respectively, were predicted by the VABS Socialization standard score, the "Interest in a familiar person" score, the "Interest in other" score and the "Distress to other" score. Both models were run a second time, replacing VABS Socialization standard score by the V-scale scores of the two subdomain scales it is composed of - the VABS Interpersonal Relationships subdomain and the VABS Play and Leisure subdomain. The logistic regressions including age and behaviour, respectively, were used additionally to account for the fact that the optima were observed to be clustered at either the parent's or stranger's face, asking whether the likelihood of converging at the parent- or the stranger-side of the stimulus space was related to age or behaviour.

Similarity between parent and stranger. Simple linear regression and simple logistic regression were used to test whether the optimum-parent Euclidean distance and the odds of converging towards parent versus stranger, respectively, depended on parent-stranger similarity.

Comparing BO results with traditional group-level results. In addition to analysing individual optima, the Nc mean amplitude was analysed on a group-level. First, an analysis was run to find out how far the group-level results of the present sample match with results from previous studies. To this end, a repeated-measures ANOVA was run of the Nc mean amplitude towards only the 100%-parent and 100%-stranger photo presented in the burn-in phase, now computed in the traditional way (mean amplitude across the 250-800 ms time window rather than for the period of negative deflection). A second analysis was run to find out how far the group-level results of the present sample match the individual-level BO optima. To this end, a repeated-measures ANOVA of the Nc mean negativity towards all presented photos of 100% parent and 100% stranger across the experiment was calculated.
Finally, the sample was split into two subsamples, once of infants who had converged closer to the parent in the stimulus space ("familiar optimum" subgroup), and once of infants who had converged closer to the stranger in the stimulus space ("unfamiliar optimum" subgroup).

5.3 Results

5.3.1 Overall attrition and convergence

Out of the 61 infants who participated, 52 infants completed the study (85%). Convergence was obtained for 44 out of the 52 infants with valid EEG data (85%), whereas the remaining 8 infants (15%) completed 15 blocks without converging to an *optimum (*Figure 5.4, top). The mean number of blocks for convergence was *m*=10.34 (SD=2.26, range: 6-14).

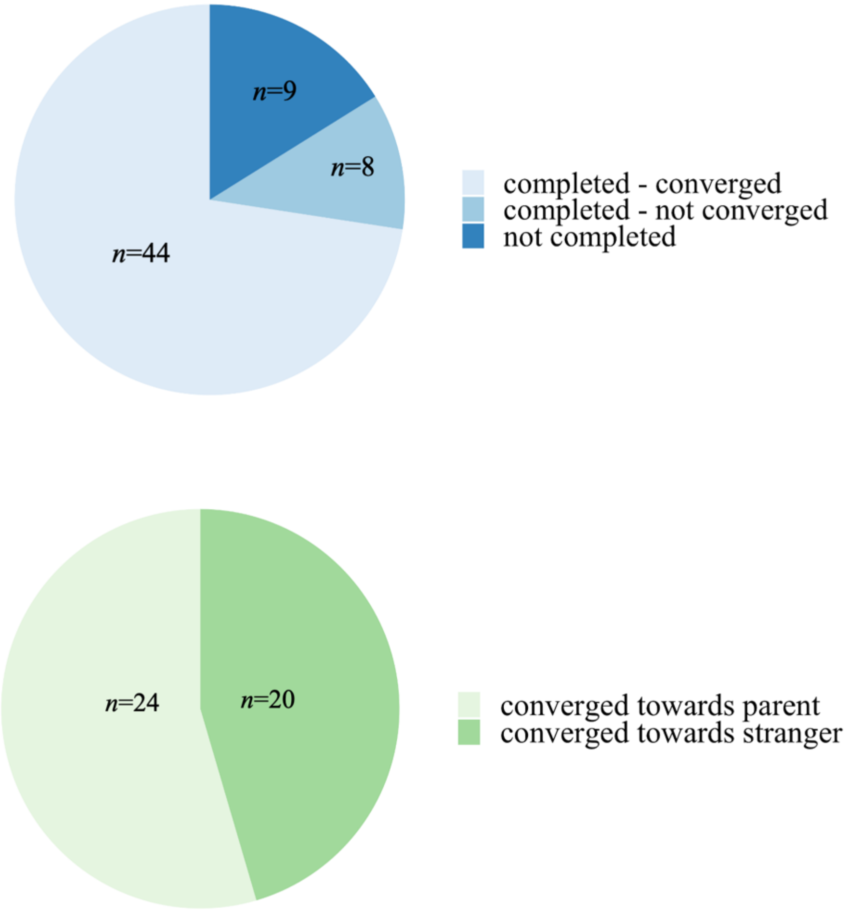


Figure 5.4. Top: Proportion of infants who completed the paradigm and converged or did not converge within 15 blocks. Bottom: Proportion of infants who converged on the parent and stranger side of the stimulus space, respectively.

5.3.2 Preference for parent versus stranger

Of the 44 infants who converged, 20 (46%) infants converged for their original parent's face, while 16 (36%) converged for the original stranger's face, and 8 infants (18%) converged for one of the morphed faces between the two extremes. Following this finding of optima not being spread across the continuous parent-stranger stimulus space, but rather clustered at either parent's or stranger's face, I split the sample into infants converging on the parent versus stranger half/side of the stimulus space. I then used additional logistic regressions to calculate effects of the variables of interest on the probability for converging towards parent versus stranger, that is on the parent versus stranger half of the stimulus space. Upon splitting the stimulus space into a parent-half and a stranger-half, of the 44 infants who converged, 24 infants (55%) converged on the parent side of the stimulus space, while 20 infants (45%) converged on the stranger side (Figure 5.4, bottom). A 2-sample test for equality of proportions indicated that this difference was not significant ($X^2(1,2) = 0.409$, $p = 0.522$).

The individual level NBO output is visualised in A5.5.1.

5.3.3 Relation with age

In the 44 infants who reached the early stopping criterion, the probability of *overall* convergence (i.e. reaching the early stopping criterion) was not related to age (logistic regression: $\beta = 0.002$, $SE = 0.007$, $p = 0.721$). The *optimum-parent Euclidean distance* did not differ significantly by age group (ANOVA: $F(1,43) = 0.273$, $p = 0.604$) and was not predicted by age in days (simple linear regression: $\beta = -0.004$, $SE = 0.002$, $p = 0.115$). The odds of converging for parent versus stranger did not significantly differ by age group (Fisher's test: $n = 52$, $p = .402$). The *odds of converging for parent versus stranger* was not related to age in days ($\beta = .010$, $SE = .006$, p=.081, $\eta_p^2 = 1.79$).

5.3.4 Relation with behaviour

The probability of overall convergence (i.e., reaching the early stopping criterion) was not related to infants' interest in familiar persons ($\beta = -1.689$, $SE = 2.204$, $p = 0.443$), interest in unfamiliar persons ($\beta = -0.930$, $SE = 1.854$, p=.616) and distress towards unfamiliar persons ($\beta = -0.136$, $SE = 0.318$, $p = 0.669$) in a multiple logistic regression.

The *optimum-parent Euclidean distance* was not significantly predicted by the Socialisation Standard Score of the VABS (p>.3), “Interest in a familiar person” score (β = 0.453, SE = 0.709, p = 0.527), “Interest in other” score (β = -0.265, SE = 0.654, p = 0.687) or the “Distress towards others” score (β = 0.132, SE = 0.152, p=0.392) in a multiple linear regression. The *probability of converging for parent versus stranger* did not significantly differ by VABS Socialisation Standard Score (p>.5), “Interest in a familiar person” score (β = -.503, SE = 1.380, p=.716), “Interest in other” score (β = -.034, SE = 1.269, p=.979) or the “Distress towards others” score (β = -.181, SE = .297, p=.541) in a multiple logistic regression. Replacing the Socialisation Standard Score by V-scale scores (Interpersonal Relationships, Play and Leisure Time) in a second run of the model did not reveal significant effects (Interpersonal Relationships: p>.9; Play and Leisure Time: p>.8).

5.3.5 Similarity between parent and stranger

The ratings of similarity between parent and stranger by the two independent researchers were significantly correlated (p=.01, r=.31). The probability of *overall* convergence was not related to the rated similarity between parent and stranger face (β = -0.001, SE = 0.025, p = 0.984). Neither the *optimum-parent Euclidean distance* (simple linear regression: p>.3) nor the odds of converging for parent versus stranger (simple logistic regression: p>.2) were significantly predicted by similarity.

5.3.6 Traditional Nc amplitude analysis

The repeated-measures ANOVA of the traditional Nc mean amplitude towards the parent- and stranger-face presented in the burn-in phase revealed no significant effect of condition (n=57, $F(1{,}55) = 0.153$, p = 0.697), age group ($F(1{,}55) = 0.221$, p = 0.640) and their interaction ($F(1{,}55) = 0.034$, p = 0.853). Adding in rated parent-stranger similarity did not change the pattern of results (all ps > 0.6).

The repeated-measures ANOVA of the Nc mean negativity towards all photos of 100% parent vs. 100% stranger presented in the course of the experiment did not significantly differ between parent and stranger (F(1,42) = 0.63, p = 0.43, η_p^2 = 0.01), age group (F(1,42) = 0.22, p = 0.64, η_p^2 = 0.01) and their interaction (F(1,42) = 0.83, p = 0.37, η_p^2= 0.02; Figure 5.5), consistent with the distribution of the individual-level optima and with the above results of the traditional Nc mean amplitude. Adding rated parent-stranger similarity as covariate to the model did not change the pattern of results, and similarity itself did not have a significant effect on the Nc ($F(1,41) = 0.09$, $p = 0.77$, $\eta_p^2 = 0.002$).

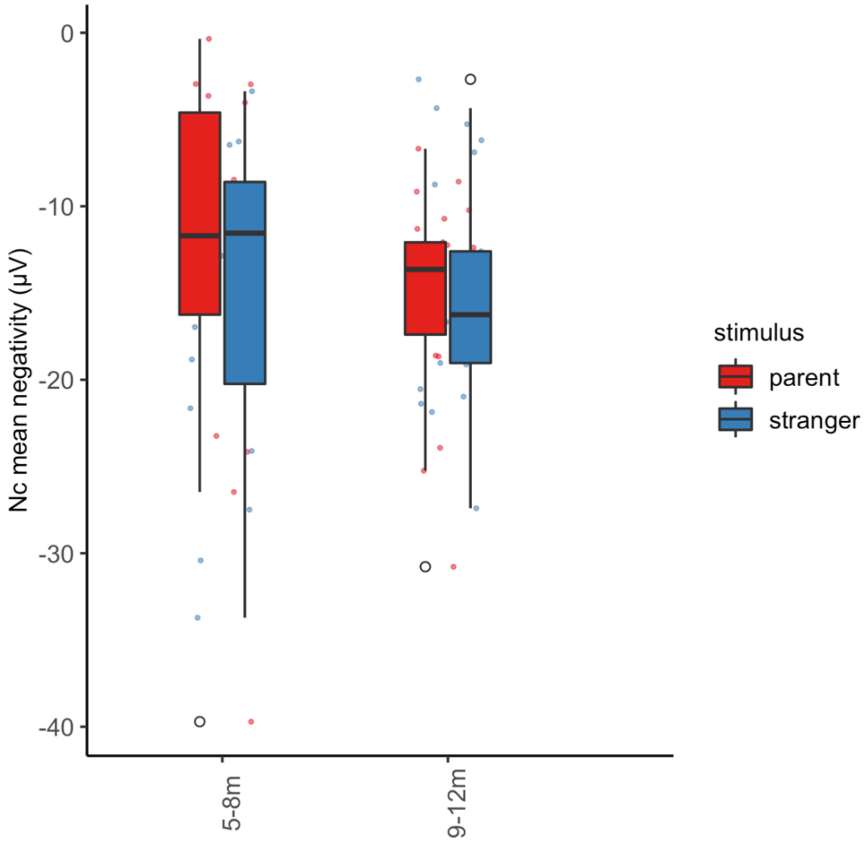
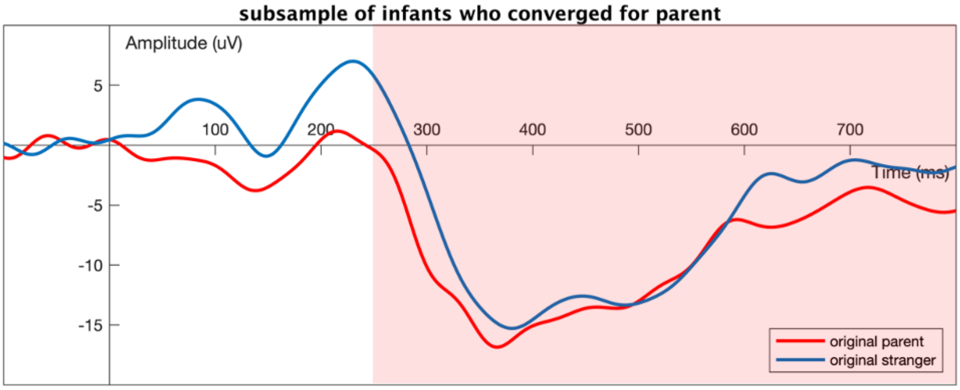


Figure 5.5. Nc mean negativity (in microvolt) by original parent and stranger face, per age group.

This analysis was re-run within each subgroup: in the “familiar optimum subgroup”, n=24, and the “unfamiliar optimum” subgroup, n=20, respectively. The Nc negativity was significantly stronger toward parent versus stranger in the “familiar optimum” subgroup ($F(1,23) = 15.39$, $p < 0.001$, $\eta_p^2 = 0.40$), and significantly stronger toward stranger versus parent in the “unfamiliar optimum” subgroup ($F(1,19) = 30.39$, $p < 0.001$, $\eta_p^2 = 0.62$; Figure 5.6).



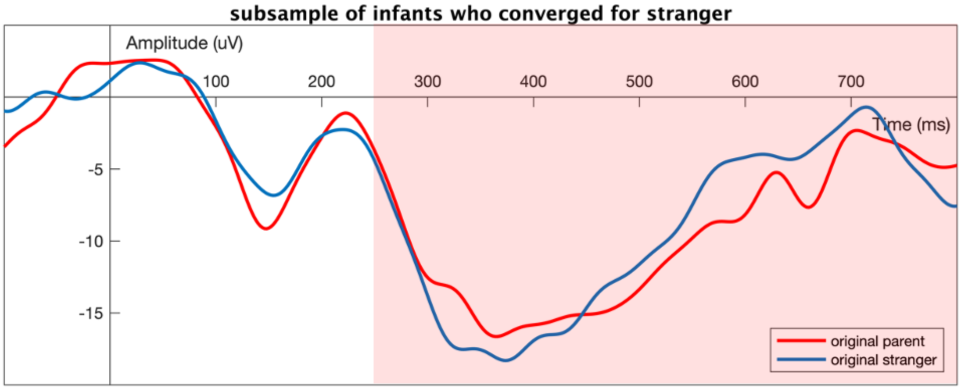


Figure 5.6. Grand average Nc waveforms across all 100% parent and 100% stranger photos across the experiment (top: subsample of infants who converged at the parent side of the stimulus space; bottom: subsample of infants who converged at the stranger side of the stimulus space)

5.4 Discussion

The present study is a proof-of-principle study combining NBO with infant ERP data to study individual infants' attention engagement with faces resembling the face of their parent and the face of a stranger. In this closed-loop experiment, the Nc mean amplitude was calculated in real time after repeated presentation of each face stimulus to an individual infant. An optimisation algorithm mapped the individual's response across various faces varying in the degree of similarity to mother's face aiming to identify the stimulus reliably producing the strongest Nc response in that infant. Identification of this optimal stimulus ("convergence") was indicated by repeated sampling of the same stimulus. The position of the individual's optimum in the stimulus space was used in offline analyses to find out whether individual optima were related to age and measures of social behaviour or depended on the rated similarity between parent and stranger.

Overall attrition rate was 15% and thus lower than the average attrition rate in traditional infant ERP studies. The BO converged to an optimum in 85% of the infants who completed the paradigm (i.e. either converged or reached the maximum of 15 blocks), that is in 72% of all infants who participated. Overall convergence did not depend on age, social behaviour measures or similarity between parent and stranger.

The proportion of infants who converged at the parent versus the stranger side of the stimulus space did not significantly differ; the probability of converging at the parent side of the space was not related to age or social behaviour measures. Group results of the Nc negativity were consistent with the individual optima results, and also with results of the Nc mean amplitude when calculated in the traditional way across the standard time-window. Further, crucially, the Nc negativity was significantly stronger for parent versus stranger in the subsample of infants who had their optimum on the parent side of the parent-stranger space, and stronger for stranger versus parent in the subsample of infants who had their optimum on the stranger side of the space, confirming that the identified optima are robust, showing individual attentional preferences for particular faces (which are not related to age or social behaviour though in the present cohort). Null-effects on the whole-group level reported in the literature in middle infancy may therefore be explained by heterogeneity in attentional preference, that is individual/subgroup preferences in different directions that cancel out each other on the group level, as opposed to intermediate or no preferences.

5.4.1 Evaluation of the proof-of-principle: NBO with infant EEG data

The approach has proved practical, given that attrition rate was lower than in classic designs, with 15% in the present study versus 21-23% in classic infant EEG studies (Van der Velde & Junge, 2020). It is possible that the reduced attrition due to infant fussiness or tiredness was related to characteristics of the BO study design, with infant-guided stimulus presentation giving them the power to "bring up" images they previously showed enhanced attention to. Also, the increased variety of stimuli compared to few, repeatedly presented stimuli might have increased infants' overall interest in and commitment to the task, hence reducing drop out.

The approach has proved reliable, given that convergence in 85% of the infants who completed the paradigm is consistent with the prediction of overall convergence of at least 50% of the sample (representing the younger infants). Convergence requires consistent empirical values, based on which the algorithm can predict a maximum in the modelled function, under low uncertainty. Convergence is a strong sign for both a preference for the identified point and for reliable data. The fact that convergence was achieved in 85% of the infants suggests that the Nc target response metric fed to the algorithm was highly consistent in that it allowed to rapidly reduce the uncertainty in the surrogate model, leading to identification of the optimum.

By contrast, inconsistent empirical values lead to an increase in uncertainty and the algorithm to continue trying to reduce the uncertainty by sampling further points, prolonging the experiment, and risk introducing more noise to the model due the infant becoming fussy with increasing duration. Reasons for non-convergence in an individual include the lack of preference between different parts of the stimulus space, and an unreliable target response metric. In both cases, the empirical values mapped onto the stimulus space do not allow to reduce uncertainty in the surrogate model. In the former case, they vary due to no preference, in the latter case, because the signal-to-noise ratio was not strong enough to reliably map the true response function.

Besides implementing strong pre-processing steps, there is no way to distinguish post-experiment whether the lack of preference or a weak signal accounted for non-convergence in an infant. While this partially also holds true in classic offline studies, here researchers are able pre-process data individually; while in the real-time case a predefined pipeline is applied to each infant, and after the experiment the researcher has only available the data recorded in response to the stimuli that were presented based on previous, potentially unreliable responses, introducing bias to an offline analysis of the raw data with different pre-processing parameters. Together, the fact that nonconvergence may imply either a lack of preference or a weak signal highlights the importance of reducing noise during recording and developing a strong pre-processing pipeline based on pilot data.

The average number of blocks needed for convergence was 10.34, corresponding to 10.34 x ~10 = 180 seconds, or ~3 minutes. Of note, in reality the experiments took longer than that on average, given that additional time was needed if blocks were repeated due to bad data quality, and for allowing the infant to bring their attention back to the screen before the start of each stimulus presentation block.

The approach has proved valid, given that the Nc negativity values were significantly stronger for parent versus stranger in the subsample of infants who converged closer toward the parent in the search space, and significantly stronger for stranger versus parent in the subsample of infants who converged closer toward the stranger in the search space. The fact that individual optima can identify subgroups of infants demonstrates that optima were indeed reflecting enhanced Nc activity and thus represent robust individual attentional preferences for particular faces. In addition, the findings from the group Nc amplitude analysis were consistent with the BO findings; both indicated no overall difference in preference for parent versus stranger, no significant relation between preference for parent versus stranger and social behaviour, and that preferences were not modulated by similarity between parent's and stranger's face, adding further support to the claim that optima represent point predicting the strongest Nc activation on the individual level. As a final note on validity, given that there is no ground truth in the literature on the pattern of Nc responses to parent/stranger in this age range, it is pointless (and also not necessary) to try to compare the present findings with the literature in order to draw conclusions about the validity of the present findings. In fact, the individual-level results of the present study can offer an explanation of the mixed pattern of findings reported previously.

5.4.2 Preference for parent versus stranger

The present study did not reveal a clear *overall* parent/stranger difference on the level of individual optima, nor the group-level Nc. However, the fact that robust optima were identified on the individual level demonstrates that this lack of an overall/group-level preference can be explained by *heterogeneity* in attentional preference, that is individual/subgroup preferences in different directions that cancel out each other on the group level, as opposed to intermediate or no preferences on the individual/subgroup level. This finding might also explain the pattern of inconsistent findings and null-effects on the whole-group level reported in the literature in middle infancy. In other words, when analysed on the group level, the effect might appear in one direction in some samples (e.g., Webb et al., 2011), while in the other direction in other samples (e.g., Swingler, Sweet & Carver, 2007), and in yet others it might not appear at all (Luyster et al., 2014). Group-level findings may depend strongly on the characteristics of the sample, as certain subgroups may show different preferences than others.

5.4.3 Relation between optima and age

In the present sample, age (both on the level of days, and when comparing 9-to-12- with 5-to-8-months-olds) did not seem to play a role in predicting whether an infant preferred more familiar or unfamiliar faces. This suggests that, contrary to what was predicted, preferences for parent or stranger do not change across the second half of the first year of age. Previous studies observed a greater Nc response for *parent* versus stranger at 6 months of age (De Haan and Nelson, 1997, 1999; Webb et al., 2005; but see Swingler, Sweet & Carver, 2007 for the reversed pattern, and Luyster et al., 2014 for no difference) and that it becomes stronger for stranger versus parent later in development (e.g., Carver et al., 2003; Dawson et al., 2002; Key & Stone, 2012; Luyster et al., 2011). The current sample, by contrast, included a broader age group (5-8 months in the younger group), and it is possible that the effect is only present during a narrower time window around 6 months. Therefore, to capture the effect of stronger Nc responses to parent versus stranger at 6 months, the age range of the younger group in the present sample might have been too broad.

Further, while previous findings had reported mixed results in the timepoint of *when* a shift from stronger Nc for parent's face towards stronger Nc for stranger's face was happening, there is the agreement that on the group-level there is an age-related shift in this direction at some point in late infancy or early childhood. In fact, while some studies reported the shift to have occurred already by the end of the first year, other studies suggested that the Nc negativity is stronger for parent versus stranger at least still in the first half of the second year (Webb et al., 2011) or until the end of the second year (Carver et al., 2003), or even throughout the entire period of infancy and toddlerhood (6-36m; Luyster et al., 2014). Therefore, paradoxically, to reveal an age-related change in the effect (from stronger Nc towards parent to stronger Nc towards stranger), the age range of the full sample (5-12 months) might have been too narrow.

5.4.4 Relation between optimum and social behaviour

In the present sample, individual differences in parent-reported social behaviour could not predict whether an infant preferred more familiar or unfamiliar faces. It was predicted that optima closer to the parent's face were related to lower socialisation skills measured by the VABS, because previous research has revealed that the Nc amplitude for parent versus stranger differed in children with autism (e.g., Webb et al., 2011). The previously described altered effects in children with autism were observed at 18-30 months and at 12-30 months in typical children (Webb et al., 2011), which is a broader and older age range than in the present study. Therefore, the age range of the present full sample might have been not broad enough to reveal significant relations between parent/stranger preference and social behaviour development.

Another possibility is that the socialisation measures used in the present study did just not capture what was modulating the effect in previous studies. Previous studies often investigated familial likelihood or diagnosis as modulating factor, and these obviously include a rich set of small-scale differences that at this young age might not be possible to be replaced by a single measure such as the VABS Socialisation score or specific indices measuring the child's behaviour towards familiar and new persons used in the present sample. Again, looking at older children, or using combined measures that have shown to relate to later autism diagnosis in infants could be more suitable. Previous research with 6-month-old infants reported that the child's behaviour during interaction with the parent is indicative of the Nc amplitude towards parent versus stranger, with increased looking for the parent during separation, Nc amplitudes towards mother became smaller and towards stranger becoming more negative. Hence, observed measures of social behaviour in interaction with the caregiver might be better able to explain the change in Nc amplitude responses towards parent versus stranger at this young age.

In addition to the social behaviour measures, the study included the administration of questions about infants' social environment. Particularly, parents were asked about the extent to which the infant was exposed to new faces in everyday life. Questions had been derived from the Infant-Toddler version of the HOME inventory (Bradley, Caldwell & Corwyn, 2003). These data are not included in the present thesis. Further analyses will include relating these measures to infants' optima in the parent-stranger-space. This data set will be particularly interesting in the context of the COVID-19 pandemic, since all infants who took part in the present study were born in the year 2021 or early in 2022, a time on a societal level characterised by more social distancing, less participation in social events, and faces outside the household covered by masks. While from summer 2021 officially all restrictions have been lifted, many individuals decided to continue following precautious measures, proposedly more so in the vulnerable time of early infancy. It would hence be interesting to see whether there was a relation between the number of new faces the child was exposed to in everyday life and their preference for familiar versus unfamiliar faces. If this is the case, then the fact that infants were less exposed to strangers during the pandemic might be one way to explain the lack of a preference for parent versus stranger in the younger infants.

5.4.5 Limitations and future directions

There are a few points to consider and further investigate in relation to the **practicability, reliability and validity** of the NBO with infant EEG.

First, regarding reliability, the high proportion of infants converging demonstrates that empirically sampled values were reliable, because once the algorithm had entered exploitation mode, it repeatedly identified them to elicit the strongest response. Of note, while convergence demonstrates that the *single-block-level values* are *reliable*, i.e. show low variability toward a given stimulus, it cannot show whether the obtained *optima/individual preferences* are *stable* across time and settings. To answer this question, the test-retest reliability of individual optima would have to be assessed by repeating the experiment with each infant, on different days. Stability of the individual optima would be indicated if the same infants converge at the same optima across visits. Future studies should include such a test, at least with a subset of the sample. Notably, this test would probably have to be conducted for each specific paradigm, as the extent to which individual preferences are stable may differ depending on the neural target metric as well as on the cognitive process under study.

Second, individual optima revealed that the vast majority of the convergers (82%) preferred one of the extremes of the stimulus space (100% parent or 100% stranger), while only 18% converged to a point in between. I did not have an a priori hypothesis about the number of infants converging at a point between the two extremes, since to my knowledge never before artificially created faces resulting from morphing parent's face into a stranger's face were used as stimuli in an infant study. One way to interpret this finding is that infants prefer faces that are either very familiar or very unfamiliar. However, other explanations need to be considered as well. One other possibility might be that this pattern is related to the selection of burn-in points; however, besides the two extremes, these included two intermediate points in the space, making this explanation less likely. Another possibility is that infants converged to the extremes because the original images were eliciting stronger responses simply because they differed somehow in features form the artificial images, which affected the Nc response. To rule out this option, a comparison would could be conducted of the low-level features between each pair of parent and stranger image that do not contribute to the high-level distinction between parent and stranger, particularly differences in luminance. Future studies should account for luminance differences a priori, by aligning the luminance between image pairs, for example using the SHINE toolbox (Willenbockel et al., 2010) implemented in MATLAB (as was done in Sandre, Freeman, Renault, Humphreys & Weinberg, 2022). Yet another explanation could be related to the way the BO was sampling the stimuli. In fact, if it was sampling the extremes of the space after the burn-in phase, this would reflect that a linear function was strongly predicted based on only the values collected in the burn-in phase, in which intermediate points had been presented, too. Also, crucially, in the present study, the optimisation algorithm was sampling also intermediate points of the space before converging to one or the other side. However, it might be worth re-investigating the sampling behaviour with a larger exploration/exploitation hyperparameter.

Third, what effect does habituation have on the method's success? The phenomenon of habituation describes attenuating in the response to a stimulus after its repeated presentation (e.g., Vecera & Johnson, 1995). In order to prevent habituated responses in the present paradigm, a maximum of three consecutive presentations of the same stimulus were possible – either when in three consecutive iterations the same stimulus was predicted to be the optimum, upon which the early stopping criterion for convergence is met, or when after two repetitions of the same block still the EEG data quality is below threshold, upon which the experiment would be terminated before completion. Could it still be that habituation builds up over the experiment, especially given that stimuli in the present paradigm were quite similar overall (e.g., 100% parent to 92% parent)? If habituation would have affected the Nc metric, with the Nc response to parent, for example, being high in blocks at the beginning of the experiment but being lower at the end of the experiment, this would have increased the uncertainty around the parent-point in the stimulus space, fostering further sampling at that point. Further attenuation of the Nc at that point would decrease the value the model predicts at that point, and hence (given high uncertainty due to inconsistent values and low predicted value) not further exploit this point. In fact, this pattern of habituation would result in not converging at all at a point in the space but instead keeping sampling to decrease uncertainty and maximise predicted values. Instead, the fact that most infants reached the early stopping criterion for convergence demonstrates that repeated sampling at a point did not prevent (e.g., through habituation) this point from being three times in the row the one that is predicted to produce the maximum value. In other words, habituation is controlled for by the fact that reaching the early stopping criterion is only possible with consistent values that are not attenuating as a consequence of habituation.

Fourth, I encountered a potential problem regarding the criterion for convergence that should be discussed as well. If the BO has entered exploitation mode, sampling at a point that it expects to elicit maximal empirical values, and upon repeated sampling of that point the *uncertainty increases at this point* (for whatever reason, e.g., due to attenuation in the values habituation, or due to low reliability in the target metric), it may still keep sampling this point, just not because it expects this point to be the optimum, but to reduce uncertainty there.

Of note, the point where the BO sampled three times in the row does not have to be identical with the predicted optimum. The predicted optimum can be extracted from the surrogate model after the session. However, this problem implies that reaching the stopping criterion does not always reflect identification of the optimum point. Instead, a true criterion reflecting identification of the optimum is lacking in the present data.

What concerns the interpretation of the present data, the fact that identified optima matched neural responses on the level of subsamples, as revealed in the statistical analysis, yields strong support to the claim that identified optima are not arbitrary points in the space but indeed represent regions (although perhaps not exact points) for which maxima are predicted.

What concerns future studies, considering alternative stopping criteria would be recommended. One possibility would be evaluating after each block whether the standard deviation of the sampled values at any of the points in the space falls below a threshold, while at the same time its predicted value is considerably higher than the predicted values in the rest of the space. Once these two conditions are met, the optimisation would stop, and that point would be considered to be the optimum. This kind of a stopping criterion would also reduce the number of blocks needed, especially the number of blocks presenting the same stimulus.

There are also points to consider and further investigate in relation to the **theoretical** implications of the present findings.

First, the whole-group age range was likely too young and narrow to reveal an age-related change in parent/stranger-preferences; and the age range of the younger age group (5-8 months) was likely too broad to establish the group-level parent-stranger-preference previously observed at 6 months.

Second, it remains subject to future investigation to find out what can explain the variance in individual preferences if not age and the behavioural measures used in the present study.

Third, regarding the behavioural measures, of note, the Vineland Interpersonal Relationships subdomain was not fully administered, including the ceiling level not being established in many infants, potentially inhibiting a relation of the EEG findings with behaviour. However, the distribution of the Vineland scores was normal and did not show a ceiling effect. Still, further behavioural measures collected should be related to the findings in order to see whether individual differences in social behaviour relate to parent-stranger preference. Further, measures of the child's social environment that might be relevant for the attention towards unfamiliar faces should be included in the analysis (e.g., HOME inventory items on exposure to new faces).

Finally, low-level image properties might have influenced the infant neural responses. As such, luminance differences between parents' and strangers' photos were not controlled. Therefore, together with the appearance of the faces on the images, the luminance of the parent image was merged with the luminance of the stranger image, potentially resulting in luminance gradients over the one-dimensional stimulus space. Future studies of this kind should consider luminance as a potential factor affecting the Nc negativity, and include a luminance control in the study design, for example by matching the parent's image to the luminance of the stranger's image using the SHINE toolbox (Willenbockel et al., 2010) in MATLAB (Sandre, Freeman, Renault, Humphreys & Weinberg, 2022).

5.4.6 Conclusion

The present study applied NBO to infant EEG and showed that the approach is practical, reliable and valid. Individual-level attentional preferences of a familiar or unfamiliar face were shown to be heterogeneous and may explain null-effects and inconsistent findings when responses are analysed across the whole group. Indeed, averaging responses of subgroups of infants with either attentional preference for parent or stranger revealed significant effects in the direction of the respective subgroup, demonstrating robust differences in preference between subgroups. Attentional preferences were not related to age or measures of social behaviour. Future work may advance criteria for the early stopping criterion and conduct test-retest checks to evaluate the stability of the individual attentional preferences.

CHAPTER 6: APPLYING INFANT NBO TO STUDY ATTENTIONAL PREFERENCES IN A NATURALISTIC SOCIAL CONTEXT

This study has been published:

Throm, E., Gui, A., Haartsen, R., da Costa, P. F., Leech, R., & Jones, E. J. (2023). Real-time monitoring of infant theta power during naturalistic social experiences. Developmental Cognitive Neuroscience, 63, 101300. https://doi.org/10.1016/j.dcn.2023.101300

6.1 Introduction

6.1.1 Identifying maximally engaging social cues in a naturalistic context

While numerous studies have reported stronger attention to social versus non-social information, increasing in strength and extend over the second half of the first year (e.g., Venema, Lowy, Earl & Webb, 2015), it remains unclear which aspects of the social world drive this increasing interest of infants in the social world. Why do infants prefer social versus non-social action? How do they choose to attend to some but not other options available? Do infants differ in these questions, and if so, how? Understanding which aspects of social interaction elicit attention in young infants is important to understand the roots of the development of the social interaction competencies infants display at the end of the first year of age (Striano, 2001).

The reasons for why the answer to this question remains unclear are many-fold. First, most studies used pre-recorded stimuli such as images of faces or vocal sounds that were presented to the infant in isolation. Studies using naturalistic paradigms have taken a step towards more ecological valid, naturalistic research by aiming to reflect the social environment the child is embedded in. Shifting towards more ecological and holistic paradigms that represent the real-world social situations infants naturally experience in daily life, with multi-modal cues embedded in complex contexts, enhances both construct validity and ecological validity. Beyond validity, live paradigms have been shown to be more powerful in eliciting brain activity specialised to social versus non-social cues (Jones, Venema, Lowy, Earl & Webb, 2015; McDonald & Perdue, 2018; Shimada & Hiraki, 2006), and it is possible that they potentiate social processing difficulties. This is particularly relevant for detecting early differences in socialisation, especially in heterogeneous groups such as groups including infants with elevated likelihood for autism, as there might be great variability in the responses between individuals.

Second, studies have primarily used the traditional experimental approach comparing responses to few pre-selected instead of multiple aspects of the social world, and averaging responses across infants in order to draw conclusions on the group level. However, in order to understand how on the level of subgroups and individuals, infants come to select from the myriad of options available, it is necessary to directly compare individual infants' responses between various aspects of social interaction.

Together, finding answers to the above questions requires studying individual infants' responses to a range of possible stimuli in a live social context. This can be done by turning around the traditional experimental approach with Neuroadaptive Bayesian Optimisation, which has proven successful in identifying individual infants' preferred stimuli among a range of possibilities (Chapter 5). The present study applies NBO in a live social paradigm to test the feasibility of this method in a live social context and to investigate which are the aspects of a naturalistic social stimulation that individual infants attend to most.

6.1.2 Evidence of infant neural responses to different aspects of live social interaction

Few studies so far directly compared multiple different aspects embedded in live social interaction. Direct gaze and infant-directed speech (IDS) have been shown to play a particularly important role in modulating infants' attention.

Direct compared to averted eye gaze has been shown in numerous studies over several decades to elicit differential responses in infants early on (Farroni, Csibra, Simion & Johnson, 2002; Farroni, Johnson, Brockbank & Simion, 2000; Farroni, Menon & Johnson, 2006). Infants have reported to use direct eye gaze as cue to recognise faces (Rigato, Menon, Johnson, & Farroni, 2011) and to follow gaze in the direction of attention of their social partner (Del Bianco, Falck-Ytter, Thorup & Gredebäck, 2018; Senju, Csibra & Johnson, 2008; Senju & Csibra, 2008; Szufnarowska, Rohlfing, Fawcett & Gredebäck, 2014). There is less research on the effect of eye gaze variations embedded in live contexts with a real social partner. On the behavioural level, averting eye gaze during live interaction accounted for a decrease in infant smiling for about 50% (Hains & Muir, 1996). In naturalistic ERP studies, the Nc amplitude reported to be greater during joint versus non-joint attention (Striano et al., 2006). Further, studies measuring infant connectivity have indicated that processing direct eye gaze is related to increased connectivity in different frequency bands. For example, alpha as well as theta band connectivity during live interaction was observed to be stronger during direct versus indirect gaze (Leong, Byrne, Clackson, Lam & Wass, 2017). Of note, EEG power (including theta) did not differ between conditions. Together, these findings suggest that eye gaze embedded in naturalistic contexts modulates neural measures of infant attention.

Infant-directed speech is the way in which adults adapt their speech when talking to a baby. Compared to adult-directed speech (ADS), it has an overall higher pitch, more variable pitch, is slower, and includes longer pauses, shorter and simpler sentences, more repetitions, and better articulation (Cooper & Aslin, 1990). It has been shown to involve larger lip movements (Green, Nip, Wilson, Mefferd & Yunusova, 2010), stronger rhythmic synchronisation and temporal regularity (Leong, Kalashnikova, Burnham & Goswami, 2017). It has been proposed to benefit language learning, by capturing infants attention (Cooper & Aslin, 1990), facilitating word segmentation (Thiessen, Hill & Saffran, 2005) and phoneme discrimination (Liu, Kuhl, & Tsao, 2003). In a recent review, it has been argued that IDS facilitates language learning because it elicits increased attention to the stimulus due to entrainment, that is time-locking, between neural oscillations and speech (Nencheva & Lew-Williams, 2022). The authors highlight individual differences in language learning, and propose that it is the interplay between behaviour, attention and the brain which defines which features of IDS are most optimal (Nencheva & Lew-Williams, 2022, p. 10). Entrainment to a stimulus might be facilitated if the stimulus fits the infants' cognitive abilities. IDS has been shown to modulate infant attention when presented as isolated, pre-recorded stimuli (for meta-analysis see Dunst, Gorman & Hamby, 2012; ManyBabiesConsortioum, 2020). For example, IDS versus ADS elicited differential ERP responses in 6-month-olds (Zangl & Mills, 2007) and stronger frontal theta power in 6-to-12-month-olds (Zhang et al., 2011), higher coherence between EEG waves and the speech envelope (Menn, Michel, Meyer, Hoehl & Männel, 2022), as well as reduced EEG coherence from frontal to other brain regions during social versus non-social attention, but only for infants from higher-income families and infants whose mothers used higher proportions of infant-directed speech (Lopera-Perez et al., 2022). There is less research on the effect of IDS versus ADS presented in a live social context. However, in one live study, 8- to 12-month-old infants showed stronger theta power responses to an infant-directed speech versus silence condition (Orekhova et al., 2006). Taken together, direct versus averted eye gaze and IDS versus ADS modulates infants' attention, while few studies have systematically examined how variations in eye gaze and vocal sounds during naturalistic experiences affect infant attention, and whether there are individual differences in preference for specific cues along these dimensions.

6.1.3 The present study

The present study for the first time used infant NBO in a live social paradigm, navigating through a rich stimulus space of live behaviours in the search for the behaviour that maximally elicits frontal theta power in the individual infant. I created a 2-dimensional stimulus space of live social behaviours to be acted out by experimenters in the lab. Dimensions were gaze direction, continuously changing from direct to 90-degrees-averted, and vocal content, continuously reducing the level of ostension, from infant-directed speech (IDS) over adult-directed speech (ADS) over neutral vocalisations to a non-vocal mechanical sound. NBO was used in order to test an individual's responses to a variety of possible social cues allowing the individual to choose which aspects of a naturalistic social situation to attend to, to disentangle theoretical predictions within one study, and to test whether there are individual differences in infant preferences to live social cues.

The study focused on an age range of 6-12 months, because theta power was shown to be stronger during social versus non-social naturalistic experiences at both 6 and 12 months, and increasing from 6 to 12 months (Jones et al., 2015). In line with previous research showing infant theta power to be stronger in a *non-live* non-interactional (i.e., videos) social versus non-social context as well as in a *live* non-interactional social versus non-social context (Jones, Venema, Lowy, Earl & Webb, 2015), the present study used non-interactional stimuli. An advantage of the novel method used in this chapter (BO) is that numerous conditions can be included in the study; hence, in addition to the live non-interactional conditions, I included one live *interactional* condition, to see whether that further adds to infants' attention to a live social partner (see section *4.2 Method*).

The primary aim of this study was to test the **feasibility of infant NBO in a naturalistic social context**. I hypothesised infant EEG data attrition rate to be lower than the averaged 25% observed in traditional studies measuring infant continuous EEG in 5- and 10-month-old infants (Van der Velde & Junge, 2020), given the greater variety of stimuli and the infant-guided stimulus selection.

The secondary aim was to see **where optima were located** in the 2-dimensional stimulus space. The theory of *infant natural pedagogy* proposes that infants pay particular attention to a social partner if behavioural cues are highly communicative and signal a sense of being addressed, thus facilitating learning of deliberately communicated information. If this is the case, infants should prefer the most communicative condition, that is direct gaze and contingent infant-directed speech.

As infants might differ in what stimuli they prefer, the third aim was to test the role of **individual differences** in parent-reported social behaviour for infants' preferences in social interaction. Which point in the space infants prefer might differ by their individual social behaviour profile. Based on findings indicating reduced elevation of theta power towards social versus non-social stimuli in individuals with later behavioural difficulties (e.g., Jones et al., 2016), I predicted infants with higher social behaviour scores to prefer behaviours that are most directed towards the infant such as IDS/direct gaze. Further, parental negative mood in relation with social settings, particularly social discomfort, was previously found to be related to infant measures of social attention (Jones, Venema, Earl, Lowy & Webb, 2017). Therefore, I predicted infants of parents with more positive and less negative mood to be more likely to prefer the behaviours closer to the IDS/direct gaze condition. Finally, *age* might affect infants' preferences in two ways: Older infants might show a preference towards the IDS/direct gaze condition because social brain specialisation has progressed in these infants and previous studies showed stronger effects of theta power during social versus non-social attention. On the other hand, around 9 months of age, infants develop the skill of joint attention (Mundy et al., 2007), allowing them to jointly look at an object with a social partner. Older infants might hence show a preference for situations of joint attention, reflected by the conditions including averted gaze. Therefore, I calculated an additional analysis with age but did not make a directional hypothesis about its effect.

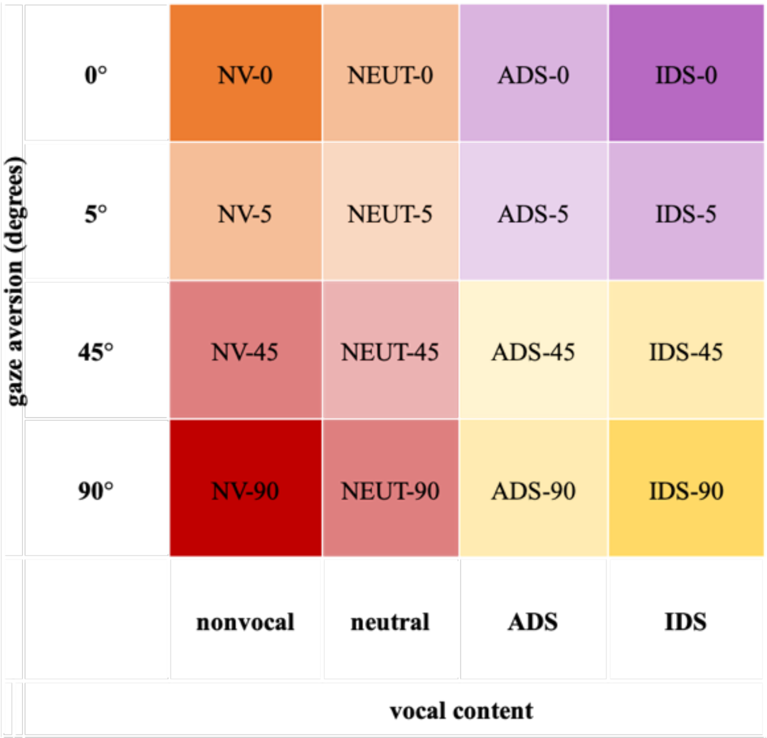
6.2 Method

6.2.1 Participants

Fifty-seven infants (20 females, 37 males) aged 6-12 months (*m*=262.60 days, *SD*=61.13 days) participated in the current study. I conducted the study in two parts, for which I split the sample. First, with n=14 infants (3 females, 11 males; age in days *m*=298.12, *SD*=49.45), the paradigm was run focussing on stimulus space exploitation, aiming to reach convergence in a minimum amount of time. Second, with n = 43 infants (17 females, 26 males; age in days *m*=251.02, *SD*=60.56), the paradigm was run focussing on stimulus space exploration, aiming to maximally explore and map out the stimulus space hence promising more detailed mapping of the individual's brain function across the stimulus space. Therefore, in this subsample I chose a highly exploratory sampling parameter so that reaching the stopping criterion was not possible.

6.2.2 Stimulus space

Stimuli were actions performed live by an actor sitting opposite the child and varied by gaze direction and vocal content. The stimulus space including 2 dimensions with 4 steps each (resulting in 16 stimuli) was created before the experiment and was equal across infants. The order of the actions along the dimensions was based on previous findings on infant theta power in response to varying gaze direction and social versus non-social sounds. The stimulus space is illustrated in Figure 6.1.



gaze aversion (degrees)	nonvocal	neutral	ADS	IDS
0°	NV-0	NEUT-0	ADS-0	IDS-0
5°	NV-5	NEUT-5	ADS-5	IDS-5
45°	NV-45	NEUT-45	ADS-45	IDS-45
90°	NV-90	NEUT-90	ADS-90	IDS-90

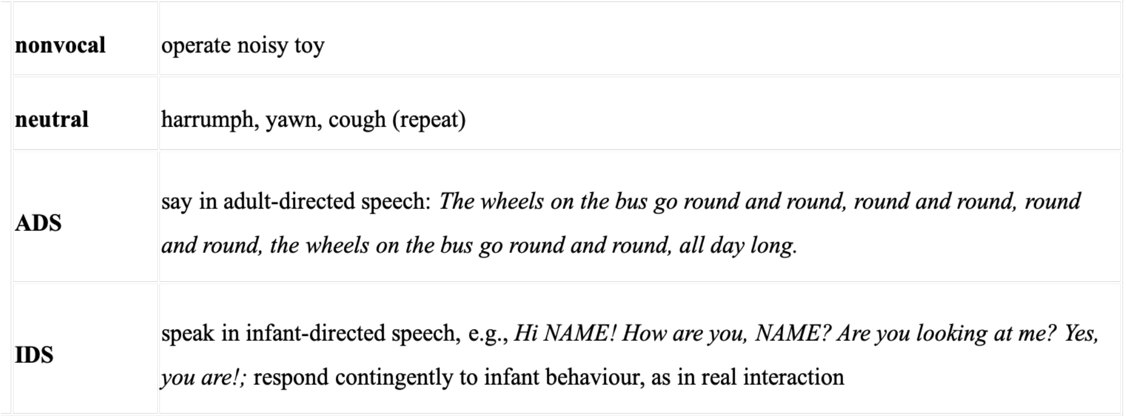
vocal content

Figure 6.1. The 2-dimensional stimulus space consists of 16 behaviours varying by gaze direction and voice. ADS = adult-directed speech, IDS = infant-directed speech

Gaze direction. Gaze direction varied from 90° averted, over 45° averted, 5° averted, to 0° averted (direct gaze). Gaze and head angle were always varied together. Previous studies showed that the gaze aversion per se is the crucial part in capturing infants' attention, while additional head aversion did not make a difference (Hains & Muir, 1996).

Vocal content. Vocal content varied from nonvocal, over vocal-neutral, adult-directed speech, to infant-directed speech. Of note, in the study by Jones and colleagues (2015) the difference in theta power between the social and non-social condition was established on the basis of merely seeing one or the other stimulus, while the auditory information was the same in both contexts, suggesting the visual input to be enough to produce a social/non-social effect. However, in order to create a powerful space that is able to elicit differential responses on the individual level and to create naturalistic stimuli as they likely may occur in the real world (i.e., direct gaze is often paired with vocalisations), the second dimension was chosen to be of auditory nature. The "nonvocal" condition (toy sound) and the "neutral vocal" condition were additionally included in reference to studies measuring social brain activation across social versus non-social sounds (Blasi et al., 2015, 2011). Conditions of the vocal content dimension are described in *Table 6.1*. Of note, the IDS condition on this dimension includes a component of contingency. However, contingency during IDS was virtually only possible during direct gaze, when the actor was not only hearing infants' vocalisations, but also seeing the infants' facial expression and bodily movement and could respond to them. On the other hand, contingency likely was decreasing with more averted gaze, due to the actor only seeing the infant in the angle of their eye. While the actor would still be able to respond to vocalisations, facial expressions and bodily movement were further out of sight.

Table 6.1. Description of the vocal content conditions presented for eight seconds per block. ADS = adult-directed speech; IDS = infant-directed speech



nonvocal	operate noisy toy
neutral	harrumph, yawn, cough (repeat)
ADS	say in adult-directed speech: *The wheels on the bus go round and round, round and round, round and round, the wheels on the bus go round and round, all day long.*
IDS	speak in infant-directed speech, e.g., *Hi NAME! How are you, NAME? Are you looking at me? Yes, you are!;* respond contingently to infant behaviour, as in real interaction

6.2.3 Actors

Different female actors (n=8) carried out the behaviours in different sessions. All actors were trained in acting out the live actions following a set procedure (*A.6.1)* to make sure all actors have a clear and the same understanding of the different behaviours in the stimulus space. Since the primary aim is the identification of the individual child's optimum, each session can be considered as an experiment in itself, in which one baby's responses to different conditions are compared rather than multiple baby's responses to few conditions. Therefore, small differences between the actors can be tolerated as long as within one actor the behaviours are clearly distinguishable as described in the stimulus space. However, an extra check was computed to ensure that the live actor did not affect the relation between theta power and the stimulus space.

6.2.4 Setup and procedure

Actors sat on a chair 1.5 m opposite the child. The child was either sitting on parent's lap or in a highchair. In very few cases, when the child was particularly fussy, the parent held the child while standing, or the child was sitting on the floor. Throughout the experiment, the actor was sitting with their body 45° averted from the child, displaying a mildly smiling, friendly facial expression. Head and gaze angle were always varied together. Before the start of the experiment and in between stimulus presentation blocks, the actor was looking at a screen, positioned 90° averted from the experimenter and the child, respectively. Instructions were presented on that same screen. The screen was shielded by a black cover from the baby's view in order to reduce distraction.

At the beginning of a block, an instruction appeared on the screen, which was after 3 seconds followed by a bell sound, indicating the start of the action. Each action lasted 8 seconds, and the end of the 8 second episode was indicated by a second bell sound. After the action episode, the actor looked back at the screen and did not produce any sound, while the signal was being processed (~ 6 seconds) until the next instruction appeared on the screen.

The infants' visual attention to the live actor was not controlled for during the experiment because this was part of the signal of interest; that is, looking away was assumed to reflect part of the response to a less engaging stimulus. The child's behaviour was video recorded for offline coding of the child's looking at the experimenter.

6.2.5 EEG recording

EEG was recorded using the wireless gel based ENOBIO EEG system (NE Neuroelectrics, Barcelona, Spain). Data were transmitted to the acquisition software via WIFI connection. Within the 10-10 EEG coordinate system, 8 electrodes were used including 6 fronto-central electrodes of interest (Fz, FC1, FC2, C1, C2 and Cz) and two reference channels (P7 and P8; Figure 6.2). Frontocentral channels were chosen because previous studies reported a strong effect of elevated frontal theta power during social stimulations (Jones et al., 2015). CMS and DRL electrodes were placed on the infants' right mastoid using NE sticktrodes. EEG data was recorded in reference to the CMS channel and digitized at 500 Hz. Before starting the experiment, the EEG data stream displayed in the acquisition software was visually inspected to ensure good quality of the EEG signal.

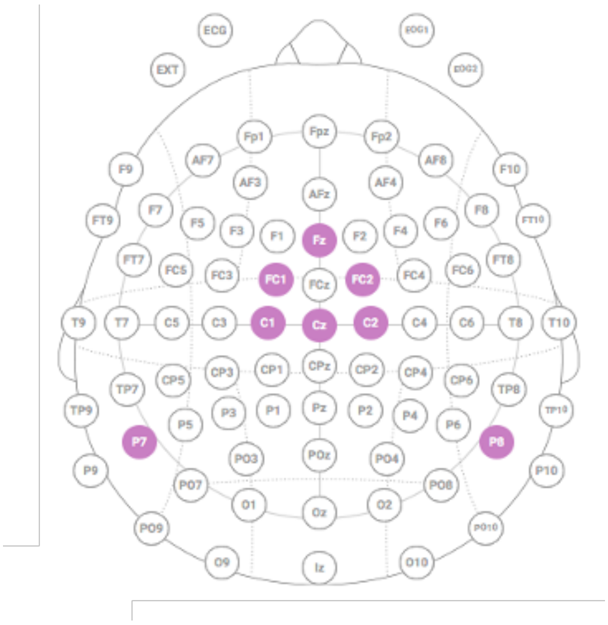


Figure 6.2. Electrode arrangement (channels of interest: Fz, FC1, FC2, C1, C2, Cz; channels for re-referencing: P7, P8)

6.2.6 Real-time processing of EEG data

The EEG data were pre-processed automatically after each stimulus presentation block using custom MATLAB scripts. Data were detrended, demeaned and band-pass filtered (0.1 to 35 Hz). The 8-second window from stimulus onset to offset was segmented from the raw data. The script confirmed that the markers were valid by checking that the start and end sample of the segment occurred after the first and before the last sample of the streamed data, respectively. The 10-second segment was cut into 1-second epochs with 50 % overlap, resulting in 19 epochs for each of the 8 channels. Given a sampling rate of 500 Hz, each 1-second epoch consisted of 500 samples of EEG data. An additional control step ensured that the last sample of the last epoch occurred before the offline marker.

For each channel in each epoch, the script checked whether any of the 500 samples exceeded an amplitude threshold of +/-200 μV or an amplitude range of 400 μV, or whether all of the 500 samples were flat (< 0.0001 mV). If at least one of these was the case, this particular channel in this particular epoch was excluded (set to NaN). Mean activation across the two re-reference channels (P7, P8) was calculated and subtracted from the raw time series of each channel of interest (Fz, Cz, FC1, C1, FC2, C2) in each of the 19 epochs (for choice of re-reference electrodes see Chapter 5).

Each channel in each epoch was subjected to a Fast Fourier Transform with a Hanning window (hannwin function, MATLAB). In each channel in each epoch, the power of the transformed signal was calculated. In each frequency band in each epoch, the time series of a channel was excluded if the power value exceeded 3 times the standard deviation of the mean of the remaining channels in that frequency band in that epoch (as in Jones et al., 2015). Data were log transformed to reduce skew. The power values in the infant theta band (3-6 Hz) was extracted and averaged to obtain absolute theta power. Relative theta power was obtained by dividing absolute theta power by the average power in all frequency bands (1-35 Hz). Relative theta power was used as key EEG metric in the present experiment.

After pre-processing, the power spectrum was plotted for each channel of interest separately as well as across all channels of interest. Further, a bar chart displayed the number of epochs in each channel that had survived artifact rejection criteria and were included in calculating the signal, allowing the experimenter to identify channels with particularly few valid epochs obtained in the current block and if needed to adjust or re-gel the electrodes to improve the EEG signal in those channels before continuing with the following block. Finally, the percentage of valid epochs across all channels of interests was displayed. If the percentage was equal to or higher than a threshold, the output was saved and used by the Bayesian Optimisation python script to select the next live behaviour from the stimulus space. If the percentage of valid time series was below the threshold, the data was not passed on to the BO and the stimulus presentation for that block was repeated.

6.2.7 Bayesian Optimisation parameters

In the present study, the Bayesian Optimisation algorithm was programmed to maximise theta power. Burn-in stimuli were the four corners of the stimulus space (nonvocal / gaze 90 averted; nonvocal / direct gaze; IDS / gaze 90 averted; IDS / direct gaze), presented to all infants in randomly assigned order.

In the exploitation sample, an exploration/exploitation hyperparameter ξ value of 0.1, which based on pilot data promised to minimise the number of blocks needed to identify the optimum in the space eliciting strongest theta power amplitude. The stopping criterion for convergence (i.e. identification of the optimum) was defined to be reached when the same point was sampled three consecutive times,that isthe point eliciting maximum theta power. If the stopping criterion for convergence was not reached, the paradigm would stop after a maximum of 15 blocks. In the exploration sample, I used a hyperparameter ξ value of 1, based on pilot data promising to allow extensive exploration of the search space, in order to reveal how the individual's brain response maps onto it.

6.2.8 Parent-report measures

The present analysis included as parent-report measures of social behaviour the *Vineland Adaptive Behaviour Scale*s v-scale scores of the subdomains "Interpersonal Relationships" and "Play and Leisure", as well as the derived scale "Distress towards Strangers" (for a description see Chapter 5).

Further, as measure of parental mood, the international short form of the *Positive And Negative Affect Scale (iPANAS-SF;* (Karim, Weisz & Rehman, 2011) was included. This scale includes 10 items, separated into measuring positive and negative affect in parents.

6.2.9 Statistical (offline) analysis

The primary outcome was the position of the individual optima in the 2D-live stimulus space obtained after each infant's session. Data from the exploitation sample were used to evaluate the BO efficiency in identifying an infant's optimum. Data from the exploitation and exploration sample were then collapsed for the remaining analyses. To do this, I made use of the fact that the algorithm updates the predicted model of the underlying function after each point of sampling, taking into account all samples collected of an infant up to that point.

In offline analyses, the relation between the position of the individual optima and individual measures (age, behaviour) was tested. The position of the individual optimum was operationalised in two ways: 1) as Euclidean distance from the most "social" condition (IDS/direct gaze) ("optimum-social distance"), with a shorter optimum-social distance reflecting that the predicted optimal stimulus was closer to the IDS/direct gaze stimulus in the 2D-stimulus space, and 2) the quadrant in which the optimum is located the stimulus space. For the latter, the 4x4-stimulus space was divided in 4 parts, including the 4 behaviours around the respective corners (Figure 6.3b).

Overall attrition and convergence. To assess attrition, the proportion of infants was calculated who did not complete the paradigm, that is for whom the experiment was terminated before reaching either convergence or the maximum number of 15 blocks, for example due to poor data quality or the infant being tired or fussy. Among the infants who completed the paradigm, the proportion of converging infants was calculated, as well as the average number of blocks needed for convergence, in order to evaluate efficacy of the algorithm.

Distribution of optima in the 2D-live stimulus space. To test the prediction of the theory of infant natural pedagogy that most optima are located in the most addressing space quadrant (Quadrant 4), I calculated the *proportion of optima per quadrant* and an ANOVA to test the *Euclidean distance between corner and optimum by quadrant*. According to the natural pedagogy theory, Quadrant 4 should show a higher proportion of optima and a smaller corner-optimum distance.

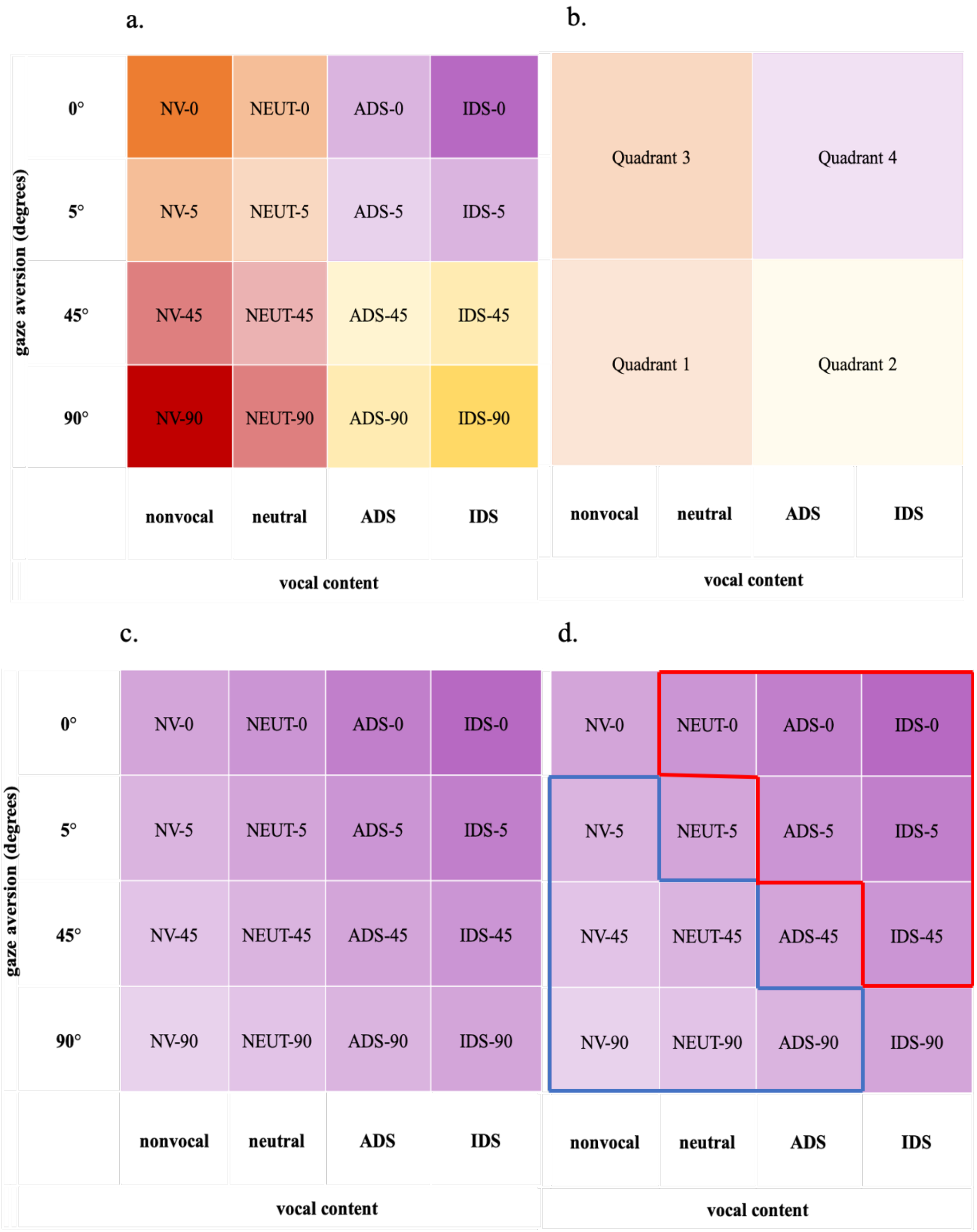
Relation with age and social behaviour. Multinomial logistic regressions were used to test whether the likelihood of converging in Quadrant 4 compared to the other Quadrants was associated with age and measures of social behaviour (VABS Interpersonal Relationships subdomain v-scale score, VABS Play and Leisure subdomain v-scale score, Distress towards Strangers score), parental mood (PANAS Positive and Negative Affect Score) and age in days. I expected that converging in Quadrant 4 was more likely in older infants and infants with higher scores on social behaviour measures, higher scores on the positive parental affect scale and lower scores on the negative parental affect scale.

Comparing BO results with traditional group-level results.

In order to compare the results produced by the BO approach with traditional group-level results within the same sample, repeated-measures ANOVA was run with theta power by quadrant.

Social/non-social structure of the stimulus space. To test how far the 2D stimulus space reflected a gradient from not socially engaging to socially engaging, spread across two dimensions of social interaction (gaze direction and vocalisation), a repeated-measures ANOVA of theta power by stimulus-space-half (socially more engaging half versus socially less engaging half; Figure 6.3d) was calculated, expecting to find greater theta power for stimuli in the socially engaging half, based on previous findings (e.g., Jones, Venema, Lowy, Earl & Webb, 2015).

Confounds by live actor. To test whether the person of the live actor was significantly modulated to the relation between theta power and stimuli, a repeated-measures ANOVA was calculated including condition (social/non-social half of the stimulus space) and live actor (8 actors).


a.

gaze aversion (degrees)	nonvocal	neutral	ADS	IDS
0°	NV-0	NEUT-0	ADS-0	IDS-0
5°	NV-5	NEUT-5	ADS-5	IDS-5
45°	NV-45	NEUT-45	ADS-45	IDS-45
90°	NV-90	NEUT-90	ADS-90	IDS-90

vocal content

b.

Quadrant 3		Quadrant 4	
Quadrant 1		Quadrant 2	
nonvocal	neutral	ADS	IDS

vocal content

c.

gaze aversion (degrees)	nonvocal	neutral	ADS	IDS
0°	NV-0	NEUT-0	ADS-0	IDS-0
5°	NV-5	NEUT-5	ADS-5	IDS-5
45°	NV-45	NEUT-45	ADS-45	IDS-45
90°	NV-90	NEUT-90	ADS-90	IDS-90

vocal content

d.

nonvocal	neutral	ADS	IDS
NV-0	NEUT-0	ADS-0	IDS-0
NV-5	NEUT-5	ADS-5	IDS-5
NV-45	NEUT-45	ADS-45	IDS-45
NV-90	NEUT-90	ADS-90	IDS-90

vocal content

Figure 6.3. Stimulus space with stimuli varying across two dimensions (y-axis: the degrees of gaze aversion, x-axis: the infant-addressed vocal content). Colours represent analytical categories of stimuli, with (a.) four extremes of the space, (b.) being divided into quadrants, c.) ranging from non-social (NV-90) to social (IDS-0), and d.) a less socially engaging half (blue) and a more socially engaging half (red).

6.3 Results

6.3.1 Overall attrition and convergence

Infants were split into two sets, one of which was tested with a low exploration/exploitation hyperparameter favouring rapid identification of the optimum stimulus (exploitative sampling), and one of which was tested with a higher exploration/exploitation hyperparameter favouring mapping the full stimulus space and hence also increasing the amount of data collected per infant (explorative sampling).

In the exploitation sample (n=14), 12 infants (86%) completed the paradigm, while in the exploration sample (n=43), 25 infants (58%) completed the paradigm. The reasons for drop-out in the exploitation sample were fussiness of the infant (n=2). The reasons for drop-out in the exploration sample were bad data quality (n=5), fussiness of the infant (n=2), an error in the script (n=6), no output saved (n=1_, or a combination of these reasons (n=4). Of note, drop-out rates were higher in the exploration sample because besides experimental failures, the experiment was longer as a consequence of the higher exploration/exploitation hyperparameter causing the algorithm to continue mapping out the space (instead of rapidly finding the optimum). The longer duration of the paradigm could be one reason for decreased data quality and increased fussiness.

In the set of infants where convergence was prioritised (exploitative sampling), 11 (92%) of the 12 infants completing the paradigm reached the early stopping criterion, and did so after on average 9 blocks (SD=2.04, range: 6-12), while one infant did not reach it and hence was presented with the maximum of 15 blocks. Since the exploration sample was explicitly programmed towards exploration, that is mapping out the stimulus space to explore the relation between theta power and the various stimuli in the search space, convergence was not reached in this sample (apart from n=1 infant).

6.3.2 Individual optima across the 2D-live stimulus space

For offline analyses, exploitation and exploration sample were collapsed to use the full dataset of infants who completed the paradigm, that is did not drop out before either reaching the early stopping criterion or the maximum of 15 blocks (n=37), making use of the fact that the algorithm updates the predicted model of the underlying function after each point of sampling, taking into account all samples collected of an infant up to that point (although optima from infants not reaching the stopping criterion are likely less robust).

An ANOVA with *Euclidean distance* between corner and individual optimum position as dependent variable, and quadrant as independent variable, revealed no significant effect of quadrant (p> .1), with median Euclidean distances of: Quadrant 1 (nonvocal/averted gaze), d=3 (mean=2.49, SD = 1.79); Quadrant 2 (IDS/averted gaze), d=3 (mean= 3.11, SD = 2.00); Quadrant 3 (nonvocal/direct gaze), d=3 (mean = 2.89, SD=2.00) from the corner of Quadrant 3; and Quadrant 4 (IDS/direct gaze), d=3 (mean=3.5, SD=1.79). In other words, according to the Euclidean distance measure, individual optima were not closer to one over other corners of the space.

Calculating the proportion of optima per quadrant, of the 37 infants who completed the paradigm, for most infants the optimum was predicted to be in Quadrant 3 (nonvocal/direct gaze, 12 or 32%), followed by Quadrant 1 and 2 (nonvocal/averted gaze and IDS/averted gaze, 10 or 27%, respectively), followed by Quadrant 4 (IDS/direct gaze, 5 or 14%). A 4-sample test for equality of proportions indicated that these proportions did not differ significantly (p > .2).

The distribution of optima visualised in *Figure 6.5* suggests that a major proportion of infants converged in one of the corners of the space, especially in the corners including nonvocal stimulation (toy). Therefore, I computed additional analyses comparing the proportion of optima located only in the corners of the space. While overall, the proportions did not differ significantly between corners in an analysis contrasting all four corners (p > .1; nonvocal/averted gaze > nonvocal/direct gaze > IDS/averted gaze > IDS/direct gaze), the proportion of optima over the two collapsed *nonvocal* corners (nonvocal/direct gaze; nonvocal/averted gaze) was significantly greater than the proportion of optima over the two collapsed *IDS* corners (IDS/direct gaze; IDS/averted gaze) (p = .049, χ^2 (1) = 3.87).

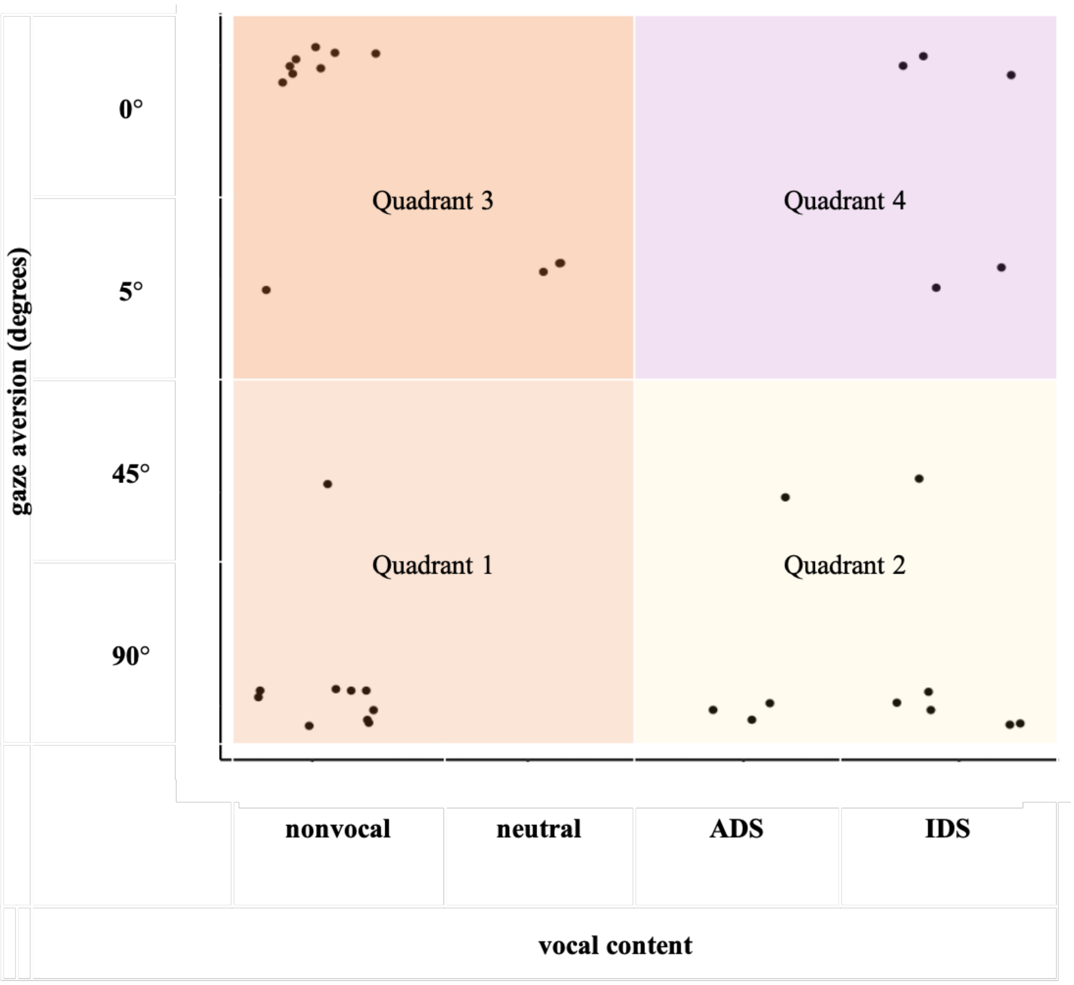


Figure 6.5. Individual optima across the 2-dimensional live interaction space of the collapsed sample (n=37).

6.3.3 Relation with age and social behaviour

Multinomial logistic regressions using Quadrant 4 (IDS/direct gaze) as reference quadrant showed that an increased likelihood of converging in Quadrant 1 (nonvocal/averted gaze) compared to Quadrant 4 (IDS/direct gaze) was significantly related to higher scores on the VABS Interpersonal Relationships subdomain (log = .55, p = .048, z = 1.97). Further, an increased likelihood of converging in Quadrant 3 (nonvocal/direct gaze; log = -2.31, p=.033, z = -2.13) or Quadrant 2 (IDS/averted gaze; log = -2.41, p=.037, z=-2.08) compared to Quadrant 4 (IDS/direct gaze) was related to lower scores on the PANAS Positive Affect subscale. No significant relations were observed with the VABS Play and Leisure subdomain (all p > .3), the derived measure "distress towards strangers" (all> .1) and age in days (all p> .2).

6.3.4 Comparing BO results with traditional group-level results

Theta power values were transformed to reduce skew and the influence of outliers (transformation method recommended by R *bestNormalize,* Peterson, 2021: orderNorm). The ANOVA of theta power by quadrant revealed no significant effect of quadrant on theta power (p = .07, F (3, 108) = 2.41, $\eta p^2 = .06$; IDS/averted gaze > nonvocal/direct gaze > nonvocal/averted gaze > IDS/direct gaze). Including age in days, Vineland Socialisation scale and parental mood in the model revealed a marginally significant effect of age (p = .08, F (1, 15) = 3.56, $\eta p^2 = .19$). Replacing the VABS Socialisation standard score by the two subdomains VABS Play and Leisure Time and VABS Interpersonal Relationships additionally revealed a marginally significant interaction effect between Quadrant and VABS Interpersonal Relationships score (p = .09, F (3, 11) = 2.79, $\eta p^2 = .02$; see A1 for a plot of the interaction).

I calculated an additional ANOVA comparing theta power across only the corners of the space. This analysis revealed a significant effect of corner on theta power (p = .028, F (3, 108) = 3.14, $\eta p^2 = .08$). Theta power was greatest in Corner 2 (IDS/averted gaze; m = 2.07, SD = 0.89), followed by Corner 3 (nonvocal/direct gaze; m = 2.04, SD = 1.99), followed by Corner 1 (nonvocal/averted gaze; m = 1.78, SD = 0.72), followed by Corner 4 (IDS/direct gaze; m= 1.23, SD = 4.59).

An additional check was computed to ensure that the person of the live actor did not affect the relation between theta power and condition. Repeated-measures ANOVA including condition (social/non-social half of the stimulus space) and live actor (8 different live actors) revealed no significant effect of the live actor in relation with condition (p = .87, F (7, 29) = 0.43, $\eta p^2 = .1$).

Social/non-social structure of the stimulus space. Theta power values were transformed to reduce skew and the influence of outliers (transformation method recommended by R *bestNormalize,* Peterson, 2021: orderNorm). A repeated-measures ANOVA of theta power by stimulus space half (social versus non-social) revealed significantly stronger theta power over the proposedly non-social versus social half (p = .018, F (1, 36) = 6.16, $\eta p^2 = .15$).

An additional repeated-measures ANOVA of theta power by only the most social (IDS/direct gaze) versus the least social condition (nonvocal/averted gaze) was calculated, revealing significantly stronger theta power in the non-social versus social condition (p < .01, F (1, 36) = 8.38, $\eta p^2 = .19$; Figure 6.4).

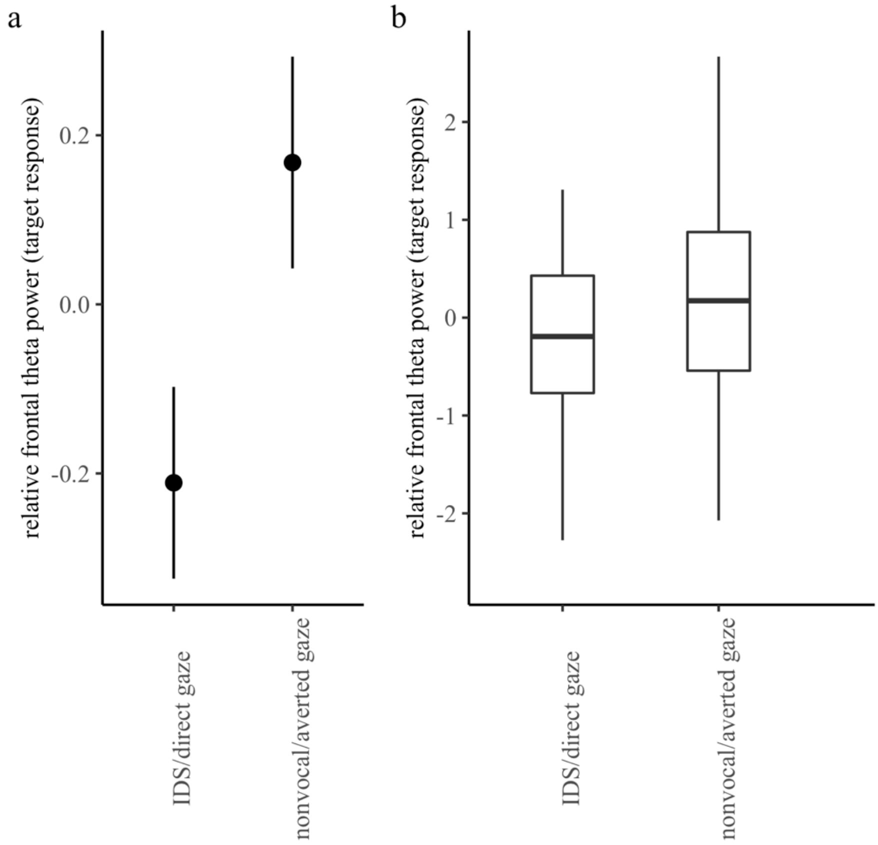


Figure 6.4. Relative frontal theta power during the most social (IDS/direct gaze) versus non-social (nonvocal/averted gaze) stimuli in the space.

6.4 Discussion

6.4.1 Summary

This study applied infant NBO to a naturalistic social paradigm to identify individual preferences of social interaction in 6- to 12-month-old infants. In specific, the BO traversed a rich 2-dimensional stimulus space of live behaviours towards identifying the behaviour that maximally elicits frontal theta power in the individual infant. Stimulus conditions varied across the stimulus space dimensions of gaze direction and vocal content directedness and were acted out by trained live actors.

In the exploitative sample, the proportion of infants who dropped out before completion of the paradigm was 14%, which is nearly half of the rate reported from traditional studies with infant continuous EEG data (25%). Of note, sacrificing exploitation in favour of broader sampling across the stimulus space came with costs on the side of attrition due to factors likely related to the increased paradigm length. In 92% of infants completing the exploitative paradigm, the BO converged, and did so after on average 9 blocks.

Infants' individual optima were not significantly more frequent in particular quadrants, although the *smallest* number of optima was observed in the IDS/direct quadrant. Instead, they were distributed across the four quadrants of the stimulus space (nonvocal/direct > nonvocal/averted and IDS/averted > IDS/direct; not significant) and were mainly clustered in the space corners (nonvocal/averted gaze > nonvocal/direct gaze > IDS/averted gaze > IDS/direct gaze; not significant).

Optima positions were related to individual measures of social behaviour and parental mood: Infants with higher VABS Interpersonal Relationships scores were more likely to prefer nonvocal/gaze averted versus IDS/direct gaze, and infants of caregivers with lower positive affect were more likely to prefer nonvocal/direct-gaze behaviours or IDS/gaze-averted behaviours compared to IDS/direct-gaze behaviours. Optima positions were not significantly related to VABS Play and Leisure scores, distress towards strangers, or age.

In traditional analyses of theta power by square across the group, theta power did not significantly differ by quadrant of the stimulus space (IDS/averted gaze > nonvocal/direct gaze > nonvocal/averted gaze > IDS/direct gaze); while they but did differ significantly by corner (IDS/averted gaze > nonvocal/direct gaze > nonvocal/averted gaze > IDS/direct gaze). Comparing only the proposedly most (IDS/direct) versus least (nonvocal/averted) social condition revealed stronger theta power for the proposedly *least* social condition.

6.4.2 Evaluating the performance of the NBO

Attrition rate in the exploitative live NBO experiment was below standard values observed in infant continuous EEG studies, likely due to the more engaging nature of the design, which allows a greater variety of stimuli and infant-guided stimulus presentation. Also, NBO experiments are shorter than traditional experiments, because it stops after enough data to predict the optimum has been collected. The low attrition rate is particularly remarkable given that this has been achieved with a live paradigm, while the standard attrition rate of 25% is based on screen-based paradigms which are normally less prone to data loss due to movement-related artifacts. Of note, attrition rate in the exploratory part of the sample was much higher, likely due to the increased paradigm length, and possibly due to the less optimum-oriented stimulus presentation causing infants to lose interest earlier.

Convergence was achieved in virtually each infant in the exploitation sample who completed the paradigm. As discussed above, convergence requires reliable empirical values, based on which the algorithm can predict a maximum in the modelled function, under low uncertainty. Unreliable empirical values lead to an increase in uncertainty and the algorithm to continue trying to reduce the uncertainty by sampling further points, prolonging the experiment, and risk introducing more noise to the model due the infant becoming fussy with increasing duration. By contrast, convergence is a strong sign for reliable empirical values. The fact that convergence was achieved in nearly all infants of the exploitation sample suggests that the EEG theta values fed to the algorithm were highly reliable in that they allowed to rapidly reduce the uncertainty in the surrogate model, leading to identification of the optimum.

The average number of blocks needed for convergence was 9, corresponding to 9 x ~ 10 = 90 seconds, or 1.5 minutes. Of note, this is the absolute minimum amount of time needed for convergence after 9 blocks; in reality, additional time is needed for blocks that have to be repeated due to bad data quality. Further, in paradigms with gaze-contingent stimulus presentation (which was not done in the current study), additional time has to be planned for allowing the child to bring their attention back to the stimulus before the start of each stimulus presentation block (e.g., Chapter 5).

Together, the present live infant NBO experiment revealed better-than-normal attrition rate when sampling in an exploitative, efficient way, and convergence after a feasible amount of time in almost all infants who completed the paradigm.

6.4.3 Which aspects of social interaction drive individual infants' attention the most?

While attentional preferences were identified in the majority of the infants, these preferences were heterogeneous and not primarily clustered in the speech/direct-gaze part of the stimulus space, contrary to what was predicted. This suggested that infant attention was not strongest towards the most direct social behaviours, contrary to what was predicted based on natural pedagogy theory. It is possible that in naturalistic contexts, individual differences in what is perceived most engaging are enhanced.

Importantly, individual optima were associated with individual differences in infants' social behaviour and environment. First, infants with stronger parent-reported social behaviour skills were more likely to prefer nonvocal/gaze averted versus IDS/direct gaze. Possibly these infants only were able to engage during the nonvocal/gaze averted behaviour with a state of joint attention. In fact, this ability has been shown to develop by the age of 9 months on a group level, and was previously shown to modulate theta power (Angelini et al., 2022; Hoehl, Michel, Reid, Parise & Striano, 2014), suggesting that infants who have already started acquiring this skill were more engaged with this condition than with others. One way to explain this finding might be that the nonvocal/gaze-averted condition has triggered some state of Joint Attention in the more socially advanced children. The ability to follow another person's gaze, the basis for engaging in Joint Attention, has been suggested to develop by 6 months of age (Gredebäck, Fikke & Melinder, 2010) and signs for Joint Attention has been observed during a live interaction paradigm in infants from 9 months of age (Cleveland & Striano, 2007). Infant theta power has shown to differ between conditions using gaze cues towards or away from objects (Angelini et al., 2022; Hoehl et al., 2014). It is possible that in the present study theta power was triggered during nonvocal/averted-gaze in infants with higher socialisation scores, because these infants were able to engage in JA and entered a state of JA when looking at the experimenter who was spinning a toy while looking 90-degrees averted.

A second relation with behaviour was observed for parental mood, with infants of parents experiencing lower positive affect being more likely to engage with the nonvocal/direct gaze or IDS/gaze averted condition, compared to the IDS/direct gaze condition. It might be that these infants show reduced attention to the IDS/direct gaze condition because previous social interactions might have been less rewarding than in other infants. In fact, parental social motivation has previously been shown to associate with infant theta power (Jones, Venema, Earl, Lowy & Webb, 2017). It then, however, remains unclear why these infants did not prefer the nonvocal/averted-gaze condition over all other conditions. Future studies are needed to investigate more in-depth infants' responses to these various aspects of live social interaction and relate them to individual differences in behaviour and social environment. Interestingly, preference was not related to age, suggesting that more than chronological age it is the developmental age and the immediate social environment that drive preference in the individuals.

While the results of the present study were not in line with the hypothesis of the theory of infant natural pedagogy proposing strongest attention to the most addressing stimulus (IDS/direct gaze), there are other theories on what aspect of social interaction drags infants' attention. Models of *contingency* or *"social expectancy"* (e.g., Murray & Trevarthen, 1986) propose contingency/reciprocity as the central factor, irrespective of other cues or combined with other cues (Hains & Muir, 1996). Contingency in the responses of a social partner refers to the temporal characteristics of the dyadic interaction, particularly the reliability and speed of feedback. This view is supported by evidence of non-contingent social stimulation from adults leading to disruptions in infant eye gaze and smiling (e.g., Bigelow, Maclean, Macdonald & Francis, 1996; ; Hains & Muir, 1996). For example, infants looked more at mother during contingent behaviour versus replayed behaviour (controlled for infant's memory, but using 1 replay where infant sees mum's replay, and 1 replay where mum sees infant's replay); and they looked more at their mother than somewhere else only during contingent but not during the replay phase (Stormark & Braarud, 2004). Finally, when mothers were presented to infants live versus delayed on TV, infants looked longer at TV live mum versus TV delayed mum; when infants were presented to mothers live versus delayed on TV, mothers did not change in their behaviour between the two; suggesting infants recognise delay (Tricia Striano, Henning, & Stahl, 2006). On the contrary, other studies failed to replicate this pattern, and thus challenge the view that non-contingency in social interaction leads to infant attention disruptions (Rochat, Neisser, & Marian, 1998). For example, 6- to 12-month-old infants did not differ in smiling and looking time between interacting with their mother live, non-live contingently, or non-live non-contingently (McClure, Chentsova-Dutton, Holochwost, Parrott & Barr, 2020), and 4-to-5-month-olds did not show stronger responses to more contingent actions per se but showed more vocalisations and smiles in response to strangers who showed a similar level of contingent responsiveness to them as their mother, instead of an overall preference for high or low contingency (Bigelow, 1998). Taken together, the major part of the evidence seems to support the hypothesis that infants are able to discriminate contingent from noncontingent conditions during live interaction, while it remains unclear which degree of contingency is most rewarding for which children. It has been proposed that while typical infants show a preference for the degree of contingency exhibited in parental responses, children with autism prefer invariable, perfectly contingent feedback, which is in real life more often found in non-social stimulation, resulting in reduced attention to social stimuli (Dawson et al., 2002). NBO studies using various behaviours varying in the degree and kind of contingency could potentially elucidate the debate and reveal which type of contingency is preferred, and whether individuals differ in this regard. Finally, it would be interesting to combine hypotheses of different theories about live social interaction in one stimulus space, varying, for example, contingency on one dimension and the degree to which a behaviour addresses the child on the other dimension, allowing to disentangle various theoretical predictions within the same NBO study.

6.4.4 Limitations

There are several limitations to the present study. First, data were collapsed across the exploitation and exploration samples in order to maximise power. However, optima in the exploration sample may have been less robust. Given the high attrition rate in the exploration sample, infant NBO in naturalistic contexts should therefore favour exploitative sampling.

Another consideration that is always important for live interaction studies (e.g., Smith et al., 2021), but particularly in an NBO context where multiple stimuli come into play, is the standardisation of acted out behaviours. All actors went through a set training procedure. However, it might be that there were differences between actors in terms of how infants responded to their behaviours. Video footage of the live actions was collected, and an additional analysis would be to have independent researchers rate the behaviours to test whether they can be considered engaging in the same way. Although there were scripts to follow for each of the behaviours, at some points were room for naturalistic dynamics that could not be pre-scripted. For example, during the IDS / direct-gaze condition, actors were responding contingently to the infant's behaviour in the moment, in order to capture naturalistic infant-directed behaviour that infants experience in everyday life, and hence the exact choice of words used differed between infants. Further, although live actors were trained to wear a "friendly smiling" facial expression throughout the session, there are likely subtle differences in facial expression between conditions.

Further, on the dimension of vocal content, besides varying vocalisation, also an action was introduced (toy spinning). It might be that due to this newly introduced modality this condition was so different from the other conditions that it attracted infants' attention particularly strongly. For example, research on infant-directed action suggested that theta power responses are elevated during infant-directed, variable movements (Meyer, van Schaik, Poli, & Hunnius, 2022).

Furthermore, the present study did on purpose not control infants' looking at the stimulus. Therefore, it remains unknown whether during an 8-second episode of live action the infant was actually looking at the stimulus. Gaze-contingent stimulus presentation was not used in this paradigm since the initial orientation towards the stimulus was aimed to be included in the response instead of preceding it. As such, not looking at the stimulus upon the start of the block was considered as part of the natural response to that stimulus. The downside of this is that the infant's visual input is not controlled for during the stimulus presentation. For example, the infant might turn away from the stimulus in order to look at mother's face; the response to mother's face would then be confounded with the response to the stimulus. On the other hand, given that BO values showed to be reliable, it can be assumed that responses were not confounded by random noise. A way to still increase experimental control without affecting the naturalism of the actions could be a control mechanism bound to data selection (as opposed to stimulus presentation). The experimenter could monitor the child's behaviour and indicate by key press the occurrence of a blink, upon which a certain window around this moment would be discarded from the analysis.

Finally, also a dichotomous social/non-social structure of the present stimulus space was tested by traditional theta power analyses. Based on strong empirical background (e.g., Chapters 3 and 4; Dawson, Bernier, et al., 2012; Haartsen et al., 2022; Jones et al., 2016; Jones et al., 2015; Orekhova et al., 2006), theta power was expected to be greater during the proposedly more socially engaging versus less socially engaging half. However, results showed the reverse pattern, with stronger theta power during the less socially engaging conditions. While the literature reports consistent and robust findings of stronger theta power during social versus non-social processing at that age, with this paradigm that includes a broader range of stimuli, the conclusions are different. In other words, the space did not reflect a linear 2D gradient from the most social to the least social condition, and there was also no between the *most* social versus *least* social condition. This suggests that the most social and least social stimuli in the present stimulus space were not consistent with the social and non-social conditions used in previous research (e.g., Jones et al., 2015; Chapter 3 and 4). The proposedly most social and least social condition in the present study were chosen based on findings from previous live social versus non-social studies measuring infant theta power. In Jones et al., 2015, the social condition was designed as looking at the experimenter, and the non-social condition as looking at the spinning toy in the experimenter's hand, while in both conditions the experimenter kept singing, spinning and looking at the infant while the infant was free to choose what to look at. In Chapter 3 and 4, the social condition was based on looking at the experimenter singing nursery rhymes, and the non-social condition as looking at the experimenter spinning a toy in their hand. In the present study, in order to create a continuum from social to non-social, singing at the extreme on the auditory dimension was replaced by infant-directed speech, because infant-directed speech was expected to be perceived more social or addressing than the other auditory stimuli (toy sound; neutral vocal sounds, that is yawning, coughing, hawking; adult-directed speech), while for singing this prediction could not be made. Further, gaze directed to the toy was replaced by 90-degrees-averted gaze, in order to allow for a continuum of gaze aversion, compared to the dichotomous categories of direct or averted. Hence, the most "social" condition in the present study was IDS/direct gaze compared to singing/direct gaze in the previous studies, and the least "social" condition was toy-sound/90-degrees-averted gaze compared to singing/gaze-directed-towards plain toy (Jones et al., 2015) and toy sound/gaze-directed-towards spinning toy (Chapter 3 and 4).

Further, in the previous studies, only segments while the infants were looking at the actor were included in the EEG analysis. In the present real-time study, looking was not controlled for and instead, EEG was analysed over the entire 8-seconds-time window of stimulus presentation, irrespective of the child's gaze. However, the experiment was video-recorded, and it would be important to code looking at the experimenter post-hoc, to see whether results changed when only EEG segments during which the child was looking at the experimenter were included in the analysis, as was the case in the previous studies (Jones et al., 2015; Chapter 3 and 4). The rationale of not including looking control in the present live study was a) not to compromise the naturalistic character of the study, and b) not to include part of what was the measure of interest (i.e. social attention) by discarding the segments during which the infants were not looking at the experimenter, possibly reflecting attentional disengagement. Hence, the difference in group-level findings of the present study might also reflect the fact that theta power is only elevated during social versus non-social conditions in which the infants are actively looking at the stimulus, highlighting the importance of the visual input. In fact, this thought is supported by the fact that in Jones et al., 2015, the effect was solely based on the visual input. The auditory input, by contrast, was not varied during the experiment, in which the singing was also present during the non-social condition. If, however, coding of looking behaviour and the discarding of non-looking-segments changed results of the group-level analysis, then future real-time studies should make efforts to find a way to control for looking behaviour during the experiment without compromising on the naturalistic quality of the stimuli.

6.4.5 Future directions

It would be worth developing an alternative measure reflecting credibility of the optima for the infants who did not reach the early stopping criterion in the present study; see also Chapter 5 for a discussion about alternative stopping criterion / criterion for considering the optimum identified.

Future research into further validating the BO approach may also compare results of exploitative and exploratory sampling in the same infant, in order to explore a) whether the rapidly identified optimum is indeed the “true” optimum, and b) how much further exploration can be done before introducing new uncertainty due to unreliable values. Further, future research may investigate not only identified optima, but also the entire response function across the space, to maximise the use of the exploratory sampling approach. If merely the identification of the maximum is of interest, exploitation should be prioritised over exploration in order to reduce attrition.

Future BO studies may further consider combining different metrics in one target metric. These could include further neural measures that have been shown to reflect social attention engagement in the first year of age (e.g., theta connectivity; van der Velde, White, & Kemner, 2021), as well as components of infant behaviour, such as smiling, has been suggested to be a good indicator for preference in very young infants (Muir & Hains, 1993). The more differentiating and reliable the (combined) target metric, the fewer iterations are needed for convergence.

One step to advance the infant NBO paradigms would be to look at responses in an actually interactive paradigm with infants' caregiver, in order to get as close as possible to studying infants in social interaction as it unfolds in real-life. I used only one condition that included contingency/reciprocity and thus an *inter*actional aspect. Normally, however, real social interaction always involves contingency, and it is this context in which I ultimately seek to study infants' attention. Studying infant-parent dyads during normal interaction and evaluate which behaviours infants prefer from a myriad of options offered by the parent could also be a step towards applying NBO as method in clinical contexts, for example to identify the behaviours that most strongly trigger social brain activation in infants with elevated likelihood for autism. These behaviours could then be used in parent-based interventions aimed at supporting social brain development over the first year of age, a period in which the brain already shows alterations neural processing of social information in infants with elevated likelihood for autism but at the same time is still very plastic and undergoing great changes, providing opportunity for intervention and supporting social development. As a more proximate next step, it would be important to study further aspects of live social interaction in a controlled live setting, such as contingency in adult responses to infant behaviour. Some theorists suggest that it is maximum contingency that infants prefer (e.g., Trevarthen, 1985), but there is evidence that it is the level of contingency experienced in daily interaction with their caregiver that determines their preferred level of contingency (Bigelow, 1998).

Future group studies may further look into the relations between attentional preferences and behavioural and environmental factors, and future NBO studies may include more distinct live stimuli potentially associated with joint attention in different subgroups, to study which types of behaviours are needed to trigger joint attention the most in these individuals.

Finally, it would be interesting to take into account the temporal dynamics of social interaction such as whether the behaviour presented in the previous block influenced responses to the present block.

6.4.6 Conclusion

This study used NBO with infant EEG to examine what aspect of naturalistic social maximises elevated theta responses in individual infants. The optimisation algorithm converged for most infants completing the paradigm in which convergence was prioritised over extensively mapping out the stimulus space, showing that this application of the method produced a robust signal reliably differentiating between different points in the stimulus space. Attentional preferences for particular behaviours were heterogeneous and depended on social behaviour skills and parental mood, suggesting that individual differences play a role in which aspect of social interaction is experienced most engaging. This study further demonstrates the particular utility of NBO for testing multiple stimulus dimensions within the same experiment.

CHAPTER 7: GENERAL DISCUSSION

7.1 Summary of the present findings

The overall aims of this thesis were to investigate whether and how neural activity during naturalistic social experiences differs in infants with an increased likelihood for autism, and to develop a method to identify in the individual infant from a range of possibilities the social cues that maximise neural correlates of social attention.

Autism is characterised by social behaviour difficulties first appearing in toddlerhood, which might be the result of early processing differences with cascading effects in social development. Under this developmental perspective, it is important to identify reliable early markers of later behavioural outcomes, both to understand how behaviour maps onto brain function over development in both typical and atypical development, and to provide interventions operating at an early stage of atypical developmental cascades, before behavioural symptoms start to emerge. The infant sibling design has been used to study the origin of autism by following up infants at elevated familial likelihood from early infancy through the timepoint of diagnosis at toddler age, aiming to identify early markers that predict later outcome. One measure that is altered in autism and has been suggested to reflect attention engagement is theta power during social compared to non-social stimulation. In typical children, theta power is stronger during social versus non-social stimulation, with an increase in strength and extent of this effect over the second half of the first year of age. It has been suggested that differential theta power responses to social versus non-social stimuli might be altered in the in early autism. One aim of this thesis was to investigate the role of differential theta power during naturalistic social experiences in early autism. To address restrictions in the questions we can ask scientifically arising from properties inherent to the traditional experimental approach, including pre-selection of few isolated stimuli for investigation and the requirement to average responses across the sample to cancel out individual variation, novel individualised methods are needed complementing the traditional group approach. This allows new types of questions being asked, for example in addition to how infants with familial likelihood for autism differ in brain responses, what type of stimulation this individual would need to optimally trigger a given neural activity. Further, such a method would allow studying responses towards a multitude of stimuli simultaneously, with the infant guiding which subset of possible stimuli are actually presented. This allows broader hypotheses, for example about what aspects of social interaction drive infant attention engagement most. The second aim of this thesis relates to introducing a method that maps the brain response of an individual infant across a range of stimuli, and to use this method to identify which social cues they attend to most.

Chapter 1 of this thesis covered theory, measurement techniques and evidence of infant attention, particularly highlighting the importance of studying infant social attention in naturalistic settings. Autism is a condition characterised by social interaction difficulties that might have its roots in early infancy, before symptoms become behaviourally visible. While autism can be studied on a group level, identifying alterations in brain responses to typically preferred social cues, given the great heterogeneity within this subgroup it is worth complementing the traditional approach with new strength-focused techniques allowing to ask what are the cues that would be more preferable for an individual.

Chapter 2 described the method of EEG and a recently developed experimental approach to the field of developmental neuroscience, Neuroadaptive Bayesian Optimisation (NBO). NBO has previously been shown to reliably map an individual adult's responses to various tasks within one session. NBO has particular value for neurodevelopmental research to overcome limits of the classic experimental design when studying individual differences in infant brain function, due to its ability to record reliable brain responses of an individual infant across multiple stimuli within one session. A core part of the chapter described how the present work combined NBO with real-time infant EEG, which steps were taken in creating and evaluating the pre-processing pipeline and which and how parameters were chosen for the Bayesian Optimisation, while considering the special requirements inherent of infant testing.

Chapter 3 reviewed literature showing the role of elevated theta power as a correlate of information coding, attention engagement and learning, including social attention, and indicating that theta power modulation might play a role in early autism. This question was investigated in a study 14-month-old infants with typical and elevated likelihood for autism by measuring theta power activity in a live social versus non-social context. Results replicated the typical effect in the overall sample which appeared in all cortical regions, but strongest posteriorly. Further, results showed that infants with elevated likelihood showed a reduced effect of elevated theta power during social versus non-social naturalistic experiences compared to infants at typical likelihood, while looking behaviour did not differ. This pattern was specific to autism likelihood, as not observed in infants with elevated ADHD likelihood. In the group of infants who went on to be diagnosed the effect tended to be there, but was not significant, likely given the small sample size of this group. Reduced looking in the social versus nonsocial condition at 14 months, but not theta power, was related to higher SRS scores at 36 months. Reduced theta power during social versus non-social attention at 14 months tended to predict higher scores on the ADOS social affect scale at 36 months, while this finding did not seem robust. ADHD traits were not predicted by socially selective looking or theta power at 14 months. Looking more in the social versus nonsocial condition at 14 months, but not theta power, partially predicted the trajectory of looking at a face, specifically more looks to a face in a screen-mediated dynamic scene at 14 and 24 months.

Chapter 4 went further back in infancy, reviewing literature suggesting a typical increase in differential theta power responses to social versus non-social cues over the second half of the first year of age, and asked whether the response of altered theta power modulation by social cues is already present at an age of 6 months, when in typical infants the effect has been observed in a live but not screen-based setting. Using the exact same paradigm as before, results replicated the effect in the overall sample over frontocentral and posterior electrodes. The effect did not differ in in infants at elevated likelihood at that age, suggesting that the alteration emerges in the second half of the first year of age. Of note, the present studies do not allow to compare the effects between the two age points. Future work should include age as a factor in a joint model to investigate the effect longitudinally. However, 6-month-old infants at elevated versus typical likelihood for autism, as well as those who went on to be diagnosed versus not diagnosed, showed reduced overall looking at the stimulus in the live setting; the latter group additionally showed increased overall theta power. Looking time in the social versus non-social condition at 6 months predicted the number of looks to a face at 10 and 24 months, while greater overall theta power in the live context at 6 months predicted fewer looks to a face at 10 months.

Chapter 5 presented a proof-of-principle study of NBO with infant EEG data, to study the individual's engagement, operationalised by the Nc mean amplitude, with a range of familiar and nonfamiliar faces. Previous research had suggested an initial attentional preference for parent's versus stranger's face with a subsequent change towards more attention towards stranger's face. The NBO study with n=61 infants aged 5-12 months viewing faces linearly varying in similarity to parent's face showed lower-than-normal attrition rate, and an equal proportion of infants preferably attending to a face closer towards parent and stranger, respectively. Individual preferences for parent/stranger were not related to age or social behaviour measures. Importantly, when subgrouping infants based on their individual attentional preference (closer to parent or stranger), Nc responses were stronger towards parent versus stranger in the parent-optimum-subsample, and stronger towards stranger versus parent in the stranger-optimum-subsample, demonstrating the validity of the individual optima. This study proved the practicability, reliability and validity of the NBO approach with infant neurophysiological data to successfully identify among a range of social cues the one maximally triggering a target brain response in the individual.

Chapter 6 built on the insights from Chapter 3, 4, and 4, and applied the individualised approach to a naturalistic context, investigating which aspects of naturalistic social experiences maximise theta power responses in the individual infant. Infants aged 6-12 months were viewing a live experimenter acting out behaviours varying in gaze direction and degree of vocalisation. Results showed a lower-than-normal attrition rate in the subsample of infants who took part in an exploitative (as opposed to explorative) version of the experiment. Identification of the optimum stimulus was reached in 92% of infants who completed the exploitative paradigm. There was no overall preference for particular parts in the space, challenging the natural pedagogy theory. Rather, individual preferences significantly differed by measures of social behaviour and parental mood. This final study showed the feasibility of NBO in recording reliable individual-level responses towards naturalistic stimuli, to study what is the preferred type of interaction for an individual infant.

The *present, final chapter* of this thesis discusses these findings and their limitations and integrates them into the broader literature. Directions for future research are suggested.

7.2 Theta power modulation during naturalistic social experiences in early autism

The present findings showed that theta power was stronger during social versus non-social naturalistic experiences at 14 as well as at 5 months of age. This is consistent with previous research showing stronger theta power during social versus non-social live stimulation in infancy (Bazhenova et al., 2007; Jones, Venema, Lowy, Earl & Webb, 2015; Orekhova, Stroganova & Posikera, 1999; Stroganova, Orekhova & Posikera, 1998). Infant theta power has been associated with processes of attention and learning (Begus & Bonawitz, 2020; Braithwaite, Jones, Johnson & Holmboe, 2020; Jones et al., 2020; Orekhova, Stroganova, Posikera & Elam, 2006), and infant differential theta power to social versus non-social stimulation has been related to improved social behaviour skills (Dawson, Bernier, et al., 2012; Jones et al., 2017). In line with previous research, differential theta power activation during social versus non-social live stimulation at 6 months was focused on frontal and occipital regions, and an increase in spatial extent of the effect from 6 months was observed, with the effect in the same infants expanding to all scalp regions by the end of the first year. Together, the present results support previous findings of social versus non-social theta power increasing in spatial extent towards the end of the first year.

Elevation of theta power during social live stimulation, but not looking proportion, was significantly reduced in infants with elevated likelihood for autism at 14 months. This pattern was not observed in a separate analysis of the same infants at 5 months of age (although the age effect was not directly compared longitudinally in the same model). This finding is consistent with previous research showing reduced theta power elevation in toddlers with autism (Dawson, Bernier, et al., 2012), and suggests that the alteration emerges in the second half of the first year of age. This is the time when in typical development the social brain network becomes increasingly specialised, giving rise to basic social behaviours emerging towards the end of the first year, including the basics for joint attention, understanding others' intentions and turn taking (e.g., Moll & Tomasello, 2004; Mundy, Card & Fox, 2000; Tomasello & Carpenter, 2007), and when in atypical development behavioural differences first start to emerge (Jones, Gliga, Bedford, Charman & Johnson, 2014). The altered effect was specific to infants with increased likelihood for autism, while infants at incrased likelihood for ADHD did not show differences in theta modulation by social live stimulation at both aged. This is in line with the view atyipcal *social* attention might be specific to autism, while atypicalities in domain-general attention contribute to ADHD, or to both conditions (Braithwaite et al., 2020).

While theta power elevation during social versus non-social live stimulation was reduced in 14-month-old infants with elevated autism likelihood, at the same age, theta power during social versus non-social stimulation in a screen-based context did not differ in infants with atypical development (Goodwin et al., 2021; Haartsen et al., 2022), highlighting the particular role of the naturalistic context for eliciting suble group differences. Tyipcally but not atyipcally developing infants seem to benefit from the combination of multimodal social cues in the live context. It may be that problems with processing social information are enhanced in contexts that require integration of multimodal cues within a complex and dynamic context. However, future studies are needed to directly compare the effect of the context on differential social processing in infants with typical and atyipcal development.

The reduced effect of elevated theta power during social versus non-social live stimulation in infants at elevated likelihood for autism did not differ by scalp region at both 6 and 14 months, which was consistent with findings in autistic toddlers (Dawson, Bernier, et al., 2012). Similarly, more subtle measures of the *change* in theta power revealed that only in typical toddlers, theta power over frontal and posterior regions increased with increased looking duration at social versus non-social videos, while in children with autism theta power increased looking at social versus non-social videos was associated with *decreased* theta power over posterior regions (Isaev et al., 2020), suggesting that autism-related alterations in theta responses differ between frontal and posterior regions in childhood. Future studies are needed to examine the developmental trajectory of how topographical regions are differentially affected by altered processing in emerging autism.

While there were socially-specific alterations of attention on a neural level in individuals with elevated autism likelihood at 14 months, at 6 months the present work showed domain-general alterations in these infants and those with later autism diagnosis, with reduced *overall* looking at the live stimulus; those with later autism diagnosis additionally showing increased *overall* theta power. Further, greater *overall* theta power at 6 months predicted fewer looks to a face within a dynamic scene at 10 months (albeit not the proportion of looking to the face). Furthermore, in adulthood, individuals with autism showed during a resting-state paradigm elevated *overall* theta power over frontal and prefrontal regions, as well as altered connectivity patterns (Murias, Webb, Greenson, & Dawson, 2007), suggesting that in addition to alterations specific to the social domain, domain-general differences in theta power might characterise autism throughout development. This is line with other studies reporting typical behaviour in infants with elevated likelihood for autism. Studies using the face-pop-out task have, for instance, shown typical orienting responses towards a face (Elsabbagh et al., 2013) and shorter overall, but not differential, fixation durations (Wass et al., 2015) in infants with elevated likelihood for autism. Furthermore, infants with elevated likelihood for autism performed lower on overall attention tasks at 2 and 3 months of age, without there being an effect of social stimuli in specific (Bradshaw et al., 2020). Hence, it might be that at 6 months, robust attention differences are predominantly present on a more domain-general level, including but not being restricted to social stimuli, while these overall altered attention processes might cascading effects on the one hand for the further trajectory of overall visual attention, and on the other hand specifically on *social* attention in the second half of the first year when in typical development the social brain network becomes increasingly specialised (Johnson et al., 2009).

While differences specific to *social* attention measured by theta power might not be robust enough at 6 months to distinguish between familial likelihood and outcome categories, they might well predict more continuous measures of social behaviour outcomes in later infancy and toddlerhood. In fact, a greater proportion of looking time in the social versus nonsocial condition at 6 months predicted the number of looks to a face at 10 and 24 months. This was not observed for other measures of looking at the face in toddlerhood, and not observed on the level of theta power; in contrast, theta power previously suggested to be related to longer peak look durations to a face at 10 months (Goodwin et al., 2021).

Differential theta power or looking at 6 months was not related to dimensional measures of autism at 36 months. At 14 months then, a greater proportion of looking at the live social versus non-social stimulus, albeit not theta power, was related to lower SRS scores at 36 months, and reduced theta power during social versus non-social attention was weakly associated with higher scores on the ADOS social affect scale at 36 months. Further research is needed to replicate these effects in an independent sample.

Together, the current findings support previous research observing elevated theta power during social versus non-social live stimulation at 6 and 12 months of age, and show that in infants with elevated likelihood for autism, but not ADHD, this effect is reduced at 12 months, while in a separate analysis of the same infants at 6 months it was not observed. Although the age effect requires in a next step being examined longitudinally to be able to directly compare (the size of) the age effect, this pattern of findings suggests that alterations in differential theta power responses characteristic for autism may emerge over the second half of the first year of age.

7.3 Using NBO to study individual infants' attentional preferences in social cues

The first infant NBO study presented (Chapter 5) served as proof of principle study of applying NBO for the first time to infant EEG data. This study used NBO to study the individual infant's Nc response across a range of images varying in the degree of similarity from parent's face. The Nc mean amplitude has been related to infant attention engagement (Gui et al., 2021; Guy et al., 2018; Richards, 2003; Webb et al., 2011), and has been shown to be stronger for parent's versus stranger's face around 6 months of age (De Haan & Nelson, 1997, 1999; Webb et al., 2005), while older children showed a greater Nc to strangers versus parents (Carver et al., 2003; Dawson et al., 2002). The findings as to when this direction is changing have been mixed, with some studies reporting a shift only after the first year of age (Carver et al., 2003; Sara Jane Webb et al., 2011) and some reporting a stronger response for stranger already towards the end of the first year of age (Guy et al., 2018; Key & Stone, 2012; Luyster, Wagner, Vogel-Farley, Tager-Flusberg & Nelson, 2011). The present NBO study investigated the Nc amplitude towards parent versus stranger in infants aged 6 to 12 months, using a stimulus pool of artificially created face images varying in similarity to parent's face, to find out which face was preferred by the individual infant.

The NBO approach proofed practical, reliable and valid in this proof-of-principle. Individual optima not clustered at a particular part of the space but rather were heterogeneous across the group; neither were they related to social behaviour measures and were not explained by similarity between parent and stranger. Further analyses of the remaining measures may help to find out what may explain the heterogeneity in parent/stranger-preference.

The second infant NBO study presented (Chapter 6) applied this new approach to study which aspects of live interaction individual infants prefer on the level of theta power. Stronger theta power during social versus non-social attention has been linked to attention and learning (Begus & Bonawitz, 2020; Haartsen et al., 2022; Jones, Venema, Lowy, Earl & Webb, 2015), and was shown to be reduced during naturalistic social experiences in infants at elevated likelihood for autism (Chapter 3 and 4). The present NBO study investigated individual theta power differences across two dimensions of social interaction, gaze direction and vocal auditory content, to identify the combinations of social cues that individual infants preferred in live social interaction. Results showed that there was no overall preference for a specific type of interaction, but that infants differed in which combination of gaze direction and vocal content they preferred. Specifically, infants with higher VABS Interpersonal Relationships scores being more likely to prefer nonvocal/gaze averted versus IDS/direct gaze, and infants of caregivers experiencing lower positive affect being more likely to prefer nonvocal/direct gaze or IDS/gaze averted versus IDS/direct gaze, respectively. The findings suggest that instead of an overall tendency to prefer one or the other aspect, preference depends on the stage of social development and the social environment of the individual.

The advantage of BO to sample a wide stimulus space allows broader hypotheses. In this thesis, the aim was to test the approach with infant EEG using a simple 1D space of stimuli presented on screen (Chapter 5) and a more complicated 2D space of live stimuli (Chapter 6). However, larger spaces can also be used, with more stimuli per dimension and more dimensions. This is useful for basic research, e.g. for identifying the task among various combinations of stimulus dimensions elicits a certain brain response in a given subgroup (e.g., Lorenz et al., 2018), as well as for application in the clinic, e.g., for identifying from multiple options the treatment that is most suited for a particular individual (e.g., Lorenz et al., 2021).

Of note, differences between group-level results and individual optima in an NBO experiment are not surprising. Since the group analysis and the individual analysis ask different questions, one expects them to produce different answers. While the traditional analysis looks at variations in the Nc across individuals, the BO analysis looks at how the Nc in one individual varies across stimuli. While in the former case individual differences within the group are cancelled out by averaging across all individuals in the group, in the latter case these individual differences are part of the research question. Subgrouping analyses can help obtain a measure of validity to see in how far optima represent condition-related differences in empirical values within the subsample.

7.3.1 Real-time analysis with infant EEG

The two NBO experiments presented in this thesis, one with screen-mediated and one with live interaction stimuli, both revealed a lower attrition rate of EEG data than in classic infant ERP studies (50%, Stets, Stahl & Reid, 2012; 25%, Van der Velde & Junge, 2020). While this was true for the exploitation-guided BO search, the attrition was higher for the exploration-guided BO search applied in the live interaction study.

Table 2. The number of blocks, paradigm length and attrition rate per study.

Paradigm	**Mean number of blocks**	**Mean duration of paradigm**	**Attrition rate**[3]
Screen, ERP, traditional			49.2% (Stets et al., 2012) 25% (van der Velde & Junge, 2020)
Screen, continuous EEG, traditional			25% (van der Velde & Junge, 2020)
Screen, ERP, exploitation-guided (Chapter 5)	10 (à 12 trials) + gaze-contingent presentation	115 seconds + time for re-centering attention + repeating bad-quality blocks	16%
Live, continuous EEG, exploitation-guided (Chapter 6)	9 (à 8 seconds)	162 seconds + repeating bad-quality blocks	14%
Live, continuous EEG, exploration-guided (Chapter 6)	15 (à 8 seconds)	270 seconds + repeating bad-quality blocks	42%

Further investigation could include the number of blocks that had to be repeated due to bad data quality. It is possible that the number blocks repeated due to bad data quality differed between paradigms; if in the live exploration-guided paradigm more blocks had to be repeated, due to the additional length of the paradigm, making the paradigm even longer, this could contribute to explaining for the higher attrition rate in this sample.

The lower attrition rate in the exploitative paradigms might be related to the shorter duration of the paradigm, or to the fact that exploitative sampling is more adaptive to the infants' responses in that it more rapidly shifts towards presenting stimuli that are

---

[3] Attrition rate was defined as the proportion of infants who dropped out before converging or reaching the predefined maximum number of 15 blocks.

predicted to trigger the strongest response in the infant, based on the previously sampled responses.

In both screen-based and live study, convergence was reached in nearly all infants who did not drop out before the end of the paradigm (screen: 85%; live: 92%). Convergence is defined as the timepoint of when the optimisation algorithm has found the maximum of the predicted function (Lorenz et al., 2017) and was operationalised by the algorithm selecting the same stimulus for three consecutive iterations. Hence, convergence reflects that the empirical values mapped on the stimulus space were reliable, in that sampling the same stimulus again produced a similar empirical value. By contrast, unreliable values would result in the failure to predict a maximum and reduce uncertainty, and hence the algorithm continuing to sample different points.
The high rate of convergence and inferred reliability in empirical values implies on the one hand that the target metric was meaningful in the individual infant, in that the individual response is indeed linked to the respective stimulus, because repeated sampling evokes the same response. On the other hand, reliability in empirical values reflects a high signal-to-noise ratio in the measured signal, despite the analysis being based on 10 trials of an individual infant.

Crucially, when subgrouping infants based on their optima, the pattern of brain response could identify the subgroup. That is, Nc responses were stronger towards parent versus stranger in the parent-optimum-subsample, and stronger towards stranger versus parent in the stranger-optimum-subsample. This finding demonstrates the validity of the individual optima in that it shows that they indeed reflect points at which maximal values are predicted instead of arbitrary points in the space.

Together, the present studies suggest that NBO with infant EEG data provides robust measurement of the individual's responses to a variety of screen based as well as live stimuli. There have been recent advances of new methods for artifact rejection specifically tailored to the neuroadaptive approach have been developed (Ouyang et al., 2022), which will further aid measuring a reliable signal in future NBO studies.

7.3.2 Implications of Using NBO for Infant Testing

There are several aspects that need to be considered for applying NBO to infant research. First, using NBO requires pre-scripting the analysis pipeline prior to data collection. This can be difficult in infant testing, where often there are no set standards for parameters. In comparison, in traditional paradigm, exact parameters, esp. for pre-processing, can be explored when seeing the data. Pilot data helps to identify optimal processing parameters. In general, pre-defining analysis steps is good-scientific practice across methodologies. Second, the data is being analysed each block, therefore, additional time is needed that adds to the overall length of the paradigm, which must be considered when designing the study. Third, in paradigms presenting visual stimuli on a screen, real-time monitoring of the infant's looking behaviour is necessary to allow gaze-contingent presentation. In classic experiments, looking can be controlled for after the study by coding the video footage. However, gaze-contingent stimulus presentation is an advantage for studies using visual stimuli in general, not only using real-time analysis, because it ensures that exactly as much data is collected as needed. Fourth, based on the respective block data, the script calculates real-time feedback about the data quality on the level of the block output measure. This is useful, because it allows researchers to correct for a bad signal in the moment of the data acquisition and hence reduce data loss. Fifth, stimulus presentation is guided by an individual infant's attention. This is useful because it tailors the paradigm to the infant's interest, together with the overall greater variability in stimuli promises a more engaging paradigm and potentially lower attrition rates. It is to consider that therefore, the duration of the experiment depends on the individual infant. Sixth, the paradigm is on average shorter than a classic paradigm, because the BO could converge and stop sampling further stimuli once the optimum has been identified (or at the latest after a predefined number of iterations). Finally, the success of the session depends on the quality of the signal, since the BO is not able to predict the underlying function if data points are unreliable. Hence, related to the first point, it is crucial to record data as clean as possible and use a processing pipeline that is able to effectively exclude noise in the data.

7.4 Additional limitations and future directions

Methods similar to NBO.

Infant NBO uses real-time analysis of the participant's and adaptive stimulus presentation. While the latter can be regarded as a type of reward feedback to the participant, the method is different from neurofeedback methods, in that usually through neurofeedback participants actively regulate their brain activity towards a certain target state. It is also different from brain-computer-interface research in that the latter translates brain states to technological devices.

NBO to study infant brain responses during interaction with their caregiver.

One particularly interesting future endeavour in the light of studying individuals' social attention development in naturalistic contexts would be to use NBO to study infant brain responses during interaction with their caregiver. The difference from the live interaction NBO study presented in the present thesis would that the caregiver instead of the trained experimenter would present different behaviours to the child. Researchers could prepare a set of behavioural cues varying along a dimension and train the parent in a brief training session prior to the session in acting out the behaviours. Possible dimensions among which behaviours during natural interaction could vary could be: the amount, speed and volume level of talking; the amount of looking at the child; the amount of touch used; the number of toys included in the play; the level of contingency in the response to the child's behaviour. Also, discrete variables are possible that can be ordered along a dimension, and behaviours can also be combined in one dimension (e.g., combining the amount of looking and touch in one dimension). The behavioural space must not be too complicated so that the behaviours are still feasible to be acted out without overlapping. In this scenario, the parent would receive instructions from the researchers, e.g., via earphones, about which of the trained behaviours to display next.

It has been suggested that truly natural and interactive settings are needed to study infant brain function more meaningfully (McDonald & Perdue, 2018). An even more naturalistic setting could involve parent and infant interacting freely, without prior training and instructing of the parent. A researcher video-monitoring the free play would classify observed behaviours based on their position on an a priori defined stimulus space. Upon signalling by the researcher that behaviour X was sampled, the data collected in a set time window before/around the moment of signalling would indicate which time window of the streamed data to analyse. As baseline, periods of rest could be identified at the beginning and used throughout the experiment, or the baseline would be “updated” in the course of the experiment by new periods of rest, identified by the researcher. This way, the stimulus space would not be constraint to behaviours that are not in the natural repertoire of an individual parent’s behaviours, training the parent would be avoided, and the behaviours displayed would be more naturalistic. Further, the parent would not be distracted from their child by receiving instructions from the researchers during the playing.

Limitations have been pointed out regarding studying infant responses during live interaction (Smith et al., 2021). The primary challenge is the non-standardisation that is on the one hand an advantage when aiming to study naturalistic social interaction, but on the other hand comes with difficulties in interpreting the findings. For example, classifying behaviours into categories bears the risk of large variability of behaviours within a category. This is problematic if the brain response varies between behaviours of a category, and further problematic for comparison between children, as not for all infants the same thing was measured. In order to circumvent the first problem, behavioural categories have to be ensured to be distinct enough to allow clear differences in brain responses between categories (e.g., IDS versus ADS). The second problem is automatically circumvented by NBO in that it operates on the individual level, that is it measures the responses to the various behaviours within the individual. Thus, for the interpretation of the optimum in the single child, which would be the aim of a naturalistic study with the caregiver freely interacting with the child, the crucial point is the variation of the presented behaviours, and that they are clearly distinguishable. It is less important whether parents in other dyads display the behaviour in a different way. Hence, less experimental control and more naturalistic, ecologically valid interaction would not pose problems to the interpretation of the data. Paradoxically, movement artifacts should even be decreased in a study with the caregiver as stimulus sitting opposite the child, while the child is sitting in a highchair for example.

The study presented in Chapter 6 showed that a meaningful signal can be measured from the individual infant during live interaction while the infant was sitting on parent's lap or on a highchair. Movement-artifacts likely stem from moments of the infant turning around to see the caregiver's face. If parent is sitting opposite the child, turning around would be inhibited, and the parent would act as stimulus and natural attention getter at the same time. Hence, a caregiver-based paradigm could even benefit for reducing movement artifacts in the EEG data. Finally, beyond aiding ecological validity, that is the generalisability of the findings to other situations of infant-caregiver interactions in this case, such a natural parent-driven design would increase the use of the result in a clinical, parent-mediated intervention setting, in that it seems more feasible and thus likely that parents increase the frequency of a natural behaviour of theirs compared to adopting new behaviours that are not already part of their repertoire.

Infant NBO with other techniques.

BO could also be used with other techniques to study infant social brain function, for example to identify auditory stimuli that optimally trigger social brain responses that have been shown to predict social behaviour outcomes in sleeping infants using functional resonance imaging (fMRI; Blasi et al., 2015) and in awake infants using functional near-infrared psectroscopy (Lloyd-Fox et al., 2018). Functional near infrared spectroscopy (fNIRS) is another technique that is very infant friendly and has been used in numerous studies to study social brain specialisation in the first years of age. It is less prone to movement, which is particularly relevant in a live interaction study where more movement can be expected. This allows for reduced data loss, which is particularly important in the real-time context. Compared to EEG, fNIRS provides a better spatial resolution of the activation across the cortex. If various source-detector separations are used, fNIRS also allows information about deeper cortical layers, with a depth sensitivity of about 1.5 cm. Several studies have demonstrated the feasibility of measuring meaningful infant fNIRS responses in a live social interaction context (Behrendt, Konrad, Perdue & Firk, 2020; Hakuno, Hata, Naoi, Hoshino & Minagawa, 2020; Sarah Lloyd-Fox, Széplaki-Köllod, Yin & Csibra, 2015; Nguyen, Abney, Salamander, Bertenthal & Hoehl, 2021; Piazza, Hasenfratz, Hasson & Lew-Williams, 2020; Urakawa, Takamoto, Ishikawa, Ono & Nishijo, 2015). It has further been suggested to use fNIRS to study social interaction in infants at elevated likelihood for autism (McDonald & Perdue, 2018). A future step would be to conduct a proof-of-principle study combining NBO with infant fNIRS data to study infant social attention in naturalistic paradigms. The limitations in temporal sensitivity inherent to measuring the hemodynamic response would have to be considered in the design of such a study. Specifically, the hemodynamic response takes several seconds to build up and go down again, before a meaningful target metric of the response to a stimulus could be computed. This would be challenging in an interactional context in which brain activation during contingent responding of the actor to the child is measured. However, in non-interactional live contexts as used in the present study, waiting for the response to build up and the computation to be finished and select the next action to present can occur during some unrelated baseline activity presented to the child during which the actor would not engage with the infant.

Clinical application.

Besides studying preferences in typically developing individuals, BO could be applied to clinical populations to tailor interventions to the individual's needs, allowing to explore the effect of various interventions simultaneously. For example, little is known about what it is that actually triggers social brain responses in autistic children, if less so the cues preferred by the neurotypical population. BO could help to identify which cues trigger their social brain response, either on the level of the clinical group, or, given the heterogeneity among autistic individuals, on the individual level. The target metric could be any clinically relevant brain metric that has been shown to respond differently to the typical stimuli and studying responses in infants at elevated likelihood in a free social interaction context could be particularly helpful in identifying relevant cues (Chapter 6).

Further possible applications of NBO.

For example, while the studies in the present thesis used 1- and 2-dimensional stimulus spaces, NBO studies can be designed with more stimuli per dimension, as well as with more than 2 dimensions, as BO intelligently selects the stimuli to be presented, without the need for a sufficient search. It is worth considering that larger search spaces will result in more computation time.

Moreover, it is in principle possible to apply the BO algorithm to behavioural data, possibly even in combination with hosting the study online using experiment builder platforms such as Gorilla.sc (e.g., for collecting parent-report data) or Lookit (for recording gaze).

Further, the present series of NBO studies aimed to identify the optimum stimulus in the individual infants by mapping a surrogate model of the unknown individual response function onto the stimulus space. While I only used the result of the optimum point in the space, a further result is this individual map of responses across the space. Using this map, we not only learn about individuals' one preferred stimulus, but also about their responses to all other stimuli. Individuals might prefer the same stimulus, but they might have different responses to the rest of the space. In a 2D space, two individuals might prefer the point located in the upper right corner, but one of them might have "got there" (i.e. showed medium responses for stimuli located) over the bottom right corner, and the other one over the top left corner. Future studies should take advantage of these individual maps revealing deeper insights into individual brain functions across variety of stimuli.

Crucially, longitudinal infant NBO experiments should be designed that follow up individual infants across different age points to reveal the developmental change in how their brain function maps across a stimulus space. This kind of studies would allow to study trajectories of individual preferences, towards understanding the causal mechanisms through which environmental input is processed in the brain, manifesting in behaviour.

7.5 Overall conclusion

The present thesis suggests that the elevation in theta power during naturalistic social compared to non-social attention observed in typical development is reduced in infants with familial likelihood of autism at 14 months. A reduced effect was not present in EEG data recorded at 5 months. This pattern of findings is in line with previous research showing an increase in strength and extent in differential responses to social versus non-social processing in typical infants.

The present thesis further extended the novel experimental approach of Neuroadaptive Bayesian Optimisation (NBO) to infant neurophysiological data. NBO inverts the traditional experimental approach to instead of looking at many infants' responses to few stimuli and measuring response differences in magnitude or location, to shift perspective to one infant's or a subgroup's response to many stimuli, while keeping constant the target neural measure. This enables researchers to identify stimulus preferred by an individual or subgroup, not replacing but complementing the traditional group-level approach by taking a different perspective on the same subject. In the context of emerging autism, the method adapts on the methodological level the view that neurodevelopmental conditions may be the result of meaningful, adaptive responses to altered processing in early development (Johnson, 2017). The proof-of-principle study in this thesis demonstrated the approach is practical, reliable and valid: It produced lower-than-normal attrition, a high proportion of infants converging to an individual optimum, and individual optima matching neural responses when averaged across subgroups. Future studies may implement a test-retest-analysis to evaluate the stability of individual optima across different timepoints and settings. Also, future research may develop alternative criteria for the early stopping criterion. An advanced definition for identification of the optimum point might make the paradigm even more robust and efficient.

References

Adolphs, R. (2003). Cognitive neuroscience: Cognitive neuroscience of human social behaviour. *Nature Reviews Neuroscience*, *4*(3), 165–178. https://doi.org/10.1038/nrn1056

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Amso, D., & Scerif, G. (2015). The attentive brain: insights from developmental cognitive neuroscience. *Nat Rev Neurosci.*, *16*(10), 606–619. https://doi.org/10.1038/nrn4025.The

Anderson, A. J., Perone, S., & Gartstein, M. A. (2022). Context matters: Cortical rhythms in infants across baseline and play. *Infant Behavior and Development*, *66*(June 2021), 101665. https://doi.org/10.1016/j.infbeh.2021.101665

Angelini, L., Tamburro, G., Lionetti, F., Spinelli, M., Comani, S., Zappasodi, F., … Aureli, T. (2022). Alpha and theta brain activity in 9-month-old infants during a live referential gaze paradigm. *Psychophysiology*, (September), 1–12. https://doi.org/10.1111/psyp.14198

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

Arslan, M., Warreyn, P., Dewaele, N., Wiersema, J. R., Demurie, E., & Roeyers, H. (2020). Development of neural responses to hearing their own name in infants at low and high risk for autism spectrum disorder. *Developmental Cognitive Neuroscience*, *41*(November 2019), 100739. https://doi.org/10.1016/j.dcn.2019.100739

Barry-Anwar, R., Hadley, H., & Scott, L. S. (2019). Differential neural responses to faces paired with labels versus faces paired with noise at 6- and at 9-months. *Vision Research*, *157*(September 2017), 264–273. https://doi.org/10.1016/j.visres.2018.03.002

Barry, R. J., Clarke, A. R., & Johnstone, S. J. (2003). A review of electrophysiology in attention-deficit/hyperactivity disorder: I. Qualitative and quantitative electroencephalography. *Clinical Neurophysiology*, *114*(2), 171–183. https://doi.org/10.1016/S1388-2457(02)00362-0

Barry, R. J., Clarke, A. R., Johnstone, S. J., McCarthy, R., & Selikowitz, M. (2009). Electroencephalogram θ/β Ratio and Arousal in Attention-Deficit/Hyperactivity Disorder: Evidence of Independent Processes. *Biological Psychiatry*, *66*(4), 398–401. https://doi.org/10.1016/j.biopsych.2009.04.027

Bates, D., Sarkar, D., Bates, M. D., & Matrix, L. (2007). The lme4 Package. *R Package Version 0.999375-22*. Retrieved from http://cran.r-project.org

Baxter, A. J., Brugha, T. S., Erskine, H. E., Scheurer, R. W., Vos, T., & Scott, J. G. (2015). The epidemiology and global burden of autism spectrum disorders. *Psychological Medicine*, *45*(3), 601–613. https://doi.org/10.1017/S003329171400172X

Bazhenova, O. V., Stroganova, T. A., Doussard-Roosevelt, J. A., Posikera, I. A., & Porges, S. W. (2007). Physiological responses of 5-month-old infants to smiling and blank faces. *International Journal of Psychophysiology*, *63*(1), 64–76. https://doi.org/10.1016/j.ijpsycho.2006.08.008

Bedford, R., Elsabbagh, M., Gliga, T., Pickles, A., Senju, A., Charman, T., & Johnson, M. H. (2012). Precursors to social and communication difficulties in infants at-risk for autism: Gaze following and attentional engagement. *Journal of Autism and Developmental Disorders*, *42*(10), 2208–2218. https://doi.org/10.1007/s10803-012-1450-y

Begum-Ali, J., Goodwin, A., Mason, L., Pasco, G., Charman, T., Johnson, M. H., & Jones, E. J. H. (2022). Altered theta–beta ratio in infancy associates with family history of ADHD and later ADHD-relevant temperamental traits. *Journal of Child Psychology and Psychiatry*. https://doi.org/10.1111/jcpp.13563

Begus, K., & Bonawitz, E. (2020). The rhythm of learning: Theta oscillations as an index of active learning in infancy. *Developmental Cognitive Neuroscience*, *45*, 100810. https://doi.org/10.1016/j.dcn.2020.100810

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 of the United States of America*, *113*(44), 12397–12402. https://doi.org/10.1073/pnas.1603261113

Behrendt, H. F., Konrad, K., Perdue, K. L., & Firk, C. (2020). Infant brain responses to live face-to-face interaction with their mothers: Combining functional near-infrared spectroscopy (fNIRS) with a modified still-face paradigm. *Infant Behavior and Development*, *58*(April 2019), 101410. https://doi.org/10.1016/j.infbeh.2019.101410

Bell, M. A., & Cuevas, K. (2012). EEG with Children. *Journal Cognitive Development*, *23*(1), 281–294. https://doi.org/10.1080/15248372.2012.691143.Using

Benedetto, S., Pedrotti, M., Minin, L., Baccino, T., Re, A., & Montanari, R. (2011). Driver workload and eye blink duration. *Transportation Research Part F: Traffic Psychology and Behaviour*, *14*(3), 199–208. https://doi.org/10.1016/j.trf.2010.12.001

Berger, A., & Posner, M. I. (2022). Beyond Infant's Looking: The Neural Basis for Infant Prediction Errors. *Perspectives on Psychological Science*. https://doi.org/10.1177/17456916221112918

Bigelow, A. E. (1998). Infants' sensitivity to familiar imperfect contingencies in social interaction. *Infant Behavior and Development*, *21*(1), 149–162. https://doi.org/10.1016/S0163-6383(98)90060-1

Bigelow, A. E., Maclean, B. K., Macdonald, D., & Francis, S. (1996). *Infants ' Response to Live and Replay Interactions with Self and Mother Author (s): Ann E . Bigelow , Barbara K . MacLean and Donna MacDonald Published by : Wayne State University Press Stable URL : https://www.jstor.org/stable/23087472 Infants ' Respon*. *42*(4), 596–611.

Birmingham, E., & Kingstone, A. (2009). Human social attention: A new look at past, present, and future investigations. *Annals of the New York Academy of Sciences*, *1156*, 118–140. https://doi.org/10.1111/j.1749-6632.2009.04468.x

Bjorklund, D. F., & Green, B. L. (1992). The adaptive nature of cognitive immaturity. *American Psychologist*, *47*(1), 46–54. https://doi.org/10.1037/0003-066X.47.1.46

Blasi, A., Lloyd-Fox, S., Sethna, V., Brammer, M. J., Mercure, E., Murray, L., … Johnson, M. H. (2015). Atypical processing of voice sounds in infants at risk for autism spectrum disorder. *Cortex*, *71*, 122–133.

https://doi.org/10.1016/j.cortex.2015.06.015

Blasi, A., Mercure, E., Lloyd-Fox, S., Thomson, A., Brammer, M., Sauter, D., … Murphy, D. G. M. (2011). Early specialization for voice and emotion processing in the infant brain. *Current Biology*, *21*(14), 1220–1224. https://doi.org/10.1016/j.cub.2011.06.009

Bond, C. F., & Titus, L. J. (1983). Social facilitation: A meta-analysis of 241 studies. *Psychological bulletin*, *94*(2), 265.

Bradley, R. H., Caldwell, B. M., & Corwyn, R. F. (2003). *The Child Care HOME Inventories : assessing the quality of family child care homes*. *18*, 294–309. https://doi.org/10.1016/S0885-2006(03)00041-3

Bradshaw, J., Klin, A., Evans, L., Klaiman, C., Saulnier, C., & McCracken, C. (2020). Development of attention from birth to 5 months in infants at risk for autism spectrum disorder. *Development and Psychopathology*, *32*(2), 491–501. https://doi.org/10.1017/S0954579419000233

Braithwaite, E. K., Gui, A., & Jones, E. J. H. (2020). Social attention: What is it, how can we measure it, and what can it tell us about autism and ADHD? *Progress in Brain Research*, *254*(Chapter 13).

Braithwaite, E. K., Jones, E. J. H., Johnson, M. H., & Holmboe, K. (2020). Dynamic modulation of frontal theta power predicts cognitive ability in infancy. *Developmental Cognitive Neuroscience*, *45*(July), 100818. https://doi.org/10.1016/j.dcn.2020.100818

Braukmann, R., Lloyd-Fox, S., Blasi, A., Johnson, M. H., Bekkering, H., Buitelaar, J. K., & Hunnius, S. (2018). Diminished socially selective neural processing in 5-month-old infants at high familial risk of autism. *European Journal of Neuroscience*, *47*(6), 720–728. https://doi.org/10.1111/ejn.13751

Buzsaki, G. (2006). *Rhythms of the Brain*. Oxford university press.

Carpenter, M., Akhtar, N., & Tomasello, M. (1998). Fourteen- through 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behavior and Development*, *21*(2), 315–330. https://doi.org/10.1016/S0163-6383(98)90009-1

Carver, L. J., Dawson, G., Panagiotides, H., Meltzoff, A. N., McPartland, J., Gray, J., & Munson, J. (2003). Age-related differences in neural correlates of face recognition during the toddler and preschool years. *Developmental Psychobiology*, *42*(2), 148–159. https://doi.org/10.1002/dev.10078

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, *18*(8), 414–421. https://doi.org/10.1016/j.tics.2014.04.012

Charman, T., & Baird, G. (2002). Practitioner review: Diagnosis of autism spectrum disorder in 2- and 3-year-old children. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *43*(3), 289–305. https://doi.org/10.1111/1469-7610.00022

Chawarska, K., Macari, S., & Shic, F. (2013). Decreased spontaneous attention to social scenes in 6-month-old infants later diagnosed with autism spectrum disorders. *Biological Psychiatry*, *74*(3), 195–203. https://doi.org/10.1016/j.biopsych.2012.11.022

Cheng, Y. Y., Wu, H. C., Tzeng, Y. L., Yang, M. T., Zhao, L. L., & Lee, C. Y. (2015). Feature-specific transition from positive mismatch response to mismatch negativity in early infancy: Mismatch responses to vowels and initial consonants. *International Journal of Psychophysiology*, *96*(2), 84–94. https://doi.org/10.1016/j.ijpsycho.2015.03.007

Cheour, M., Korpilahti, P., Martynova, O., & Lang, A. H. (2001). Mismatch negativity and late discriminative negativity in investigating speech perception and learning in children and infants. *Audiology and Neuro-Otology*, *6*(1), 2–11. https://doi.org/10.1159/000046804

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in Cognitive Sciences*, *16*(4), 231–239. https://doi.org/10.1016/j.tics.2012.02.007

Cleveland, A., & Striano, T. (2007). The effects of joint attention on object processing in 4- and 9-month-old infants. *Infant Behavior and Development*, *30*(3), 499–504. https://doi.org/10.1016/j.infbeh.2006.10.009

Cohen, M. X., Bour, L., Mantione, M., Figee, M., Vink, M., Tijssen, M. A. J., … Denys, D. (2012). Top-down-directed synchrony from medial frontal cortex to nucleus accumbens during reward anticipation. *Human Brain Mapping*, *33*(1), 246–252. https://doi.org/10.1002/hbm.21195

Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology*, *52*, 337–367. https://doi.org/10.1146/annurev.psych.52.1.337

Constantino, J. N., Zhang, Y., Frazier, T., Abbacchi, A. M., & Law, P. (2010). Sibling recurrence and the genetic epidemiology of autism. *American Journal of Psychiatry*, *167*(11), 1349–1356. https://doi.org/10.1176/appi.ajp.2010.09101470

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*. https://doi.org/10.1016/j.neuroimage.2020.116602

Cooper, R., & Aslin, R. N. (1990). *Preference for infant-directed speech in the first month after birth* (pp. 1584–1595). pp. 1584–1595. Child Development.

Courchesne, E., Ganz, L., & Norcia, A. M. (1981). Event-related brain potentials to human faces in infants. *Child Development*, *52*(3), 804–811. https://doi.org/10.2307/1129080

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

Csibra, Gergely, & Gergely, G. (2006). Social learning and social cognition: The role of pedagogy. *Magyar Pszichologiai Szemle*, *62*(1), 5–30. https://doi.org/10.1556/MPSzle.62.2007.1.2

da Costa, P.F., Haartsen, R., Throm, E., Mason, L., Gui, A., Leech, R., & Jones, E. J. H. (2021). Neuroadaptive electroencephalography: a proof-of-principle study in infants. *ArXiv*, 1–25. Retrieved from http://arxiv.org/abs/2106.06029

Dawson, G., Bernier, R., & Ring, R. H. (2012). Social attention: A possible early indicator of efficacy in autism clinical trials. *Journal of Neurodevelopmental Disorders*, *4*(1), 1–12. https://doi.org/10.1186/1866-1955-4-11

Dawson, G., Carver, L., Meltzoff, A. N., Panagiotides, H., McPartland, J., & Webb, S. J. (2002). Neural correlates of face and object recognition in young children with autism spectrum disorder, developmental delay, and typical development. *Child Development*, *73*(3), 700–717. https://doi.org/10.1111/1467-8624.00433

Dawson, G., Jones, E. J. H., Merkle, K., Venema, K., Lowy, R., Faja, S., … Webb, S. J. (2012). Early behavioral intervention is associated with normalized brain activity in young children with autism. *Journal of the American Academy of Child and Adolescent Psychiatry*, *51*(11), 1150–1159.

https://doi.org/10.1016/j.jaac.2012.08.018

De Haan, M. (1997). Recognition of faces and objects by 6-month-old infants (ProQuest Information & Learning; Vol. 57). Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1997-95004-124&site=ehost-live

de Haan, M., Johnson, M. H., & Halit, H. (2007). Development of face-sensitive event-related potentials during infancy. In M. de Haan (Ed.), *Infant EEG and event-related potentials*. (pp. 78–99). Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2006-20431-003&site=ehost-live

De Haan, M., Johnson, M. H., & Halit, H. (2003). Development of face-sensitive event-related potentials during infancy: A review. *International Journal of Psychophysiology*, *51*(1), 45–58. https://doi.org/10.1016/S0167-8760(03)00152-1

De Haan, M., Johnson, M. H., Maurer, D., & Perrett, D. I. (2001). Recognition of individual faces and average face prototypes by 1- and 3-month-old infants. *Cognitive Development*, *16*(2), 659–678. https://doi.org/10.1016/S0885-2014(01)00051-X

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. Retrieved from https://www.ptonline.com/articles/how-to-get-better-mfi-results

De Haan, M., & Nelson, C. A. (1999). Brain activity differentiates face and object processing in 6-month-old infants. *Developmental Psychology*, *35*(4), 1113–1121. https://doi.org/10.1037/0012-1649.35.4.1113

Dehaene-Lambertz, G. (2000). Cerebral specialization for speech and non-speech stimuli in infants. *Journal of Cognitive Neuroscience*, *12*(3), 449–460. https://doi.org/10.1162/089892900562264

Del Bianco, T., Falck-Ytter, T., Thorup, E., & Gredebäck, G. (2018). The Developmental Origins of Gaze-Following in Human Infants. *Infancy*, infa.12276. https://doi.org/10.1111/infa.12276

Di Giorgio, E., Frasnelli, E., Rosa Salva, O., Maria Luisa, S., Puopolo, M., Tosoni, D., … Vicari, S. (2016). Difference in Visual Social Predispositions between Newborns at Low-and High-risk for Autism. *Scientific Reports*, *6*(April), 1–9. https://doi.org/10.1038/srep26395

Dunst, C. J., Gorman, E., & Hamby, D. W. (2012). Preference for infant-directed speech in preverbal young children. *Center for Early Literacy Learning Reviews*, *5*(1), 1–13. https://doi.org/10.31234/osf.io/3wgxc

Ellis, C. T., Skalaban, L. J., Yates, T. S., Bejjanki, V. R., Córdova, N. I., & Turk-Browne, N. B. (2020). Re-imagining fMRI for awake behaving infants. *Nature Communications*, *11*(1), 1–12. https://doi.org/10.1038/s41467-020-18286-y

Elsabbagh, M., Fernandes, J., Jane Webb, S., Dawson, G., Charman, T., & Johnson, M. H. (2013). Disengagement of visual attention in infancy is associated with emerging autism in toddlerhood. *Biological Psychiatry*, *74*(3), 189–194. https://doi.org/10.1016/j.biopsych.2012.11.030

Elsabbagh, M., Gliga, T., Pickles, A., Hudry, K., Charman, T., & Johnson, M. H. (2013). The development of face orienting mechanisms in infants at-risk for autism. *Behavioural Brain Research*, *251*, 147–154.

https://doi.org/10.1016/j.bbr.2012.07.030

Elsabbagh, M., Mercure, E., Hudry, K., Chandler, S., Pasco, G., Charman, T., … Johnson, M. H. H. (2012). Infant Neural Sensitivity to Dynamic Eye Gaze Is Associated with Later Emerging Autism. *Current Biology*, *22*(4), 338–342. https://doi.org/https://doi.org/10.1016/j.cub.2011.12.056

Esler, A. N., Bal, V. H., Guthrie, W., Wetherby, A., Weismer, S. E., & Lord, C. (2015). The Autism Diagnostic Observation Schedule, Toddler Module: Standardized Severity Scores. *Journal of Autism and Developmental Disorders*, *45*(9), 2704–2720. https://doi.org/10.1007/s10803-015-2432-7

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(14), 9602–9605. https://doi.org/10.1073/pnas.152159999

Farroni, T., Johnson, M. H., Brockbank, M., & Simion, F. (2000). Infants' use of gaze direction to cue attention: The importance of perceived motion. *Visual Cognition*, *7*(6), 705–718. https://doi.org/10.1080/135062800501443999

Farroni, T., Menon, E., & Johnson, M. H. (2006). Factors influencing newborns' preference for faces with eye contact. *Journal of Experimental Child Psychology*, *95*(4), 298–308. https://doi.org/https://doi.org/10.1016/j.jecp.2006.08.001

Farroni, Teresa, Chiarelli, A. M., Lloyd-Fox, S., Massaccesi, S., Merla, A., Di Gangi, V., … Johnson, M. H. (2013). Infant cortex responds to other humans from shortly after birth. *Scientific Reports*, *3*, 3–7. https://doi.org/10.1038/srep02851

Farroni, Teresa, 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. https://doi.org/10.1073/PNAS.152159999

Farroni, Teresa, Johnson, M. H., & Csibra, G. (2004). Mechanisms of Eye Gaze Perception during Infancy. *Journal of Cognitive Neuroscience*, *16*(8), 1320–1326. https://doi.org/10.1162/0898929042304787

Frank, M. C., Vul, E., & Saxe, R. (2012). Measuring the Development of Social Attention Using Free-Viewing. *Infancy*, *17*(4), 355–375. https://doi.org/10.1111/j.1532-7078.2011.00086.x

Frazier, P. I. (2018). *A Tutorial on Bayesian Optimization*. (Section 5), 1–22. Retrieved from http://arxiv.org/abs/1807.02811

Fujioka, T., Mourad, N., He, C., & Trainor, L. J. (2011). Comparison of artifact correction methods for infant EEG applied to extraction of event-related potential signals. *Clinical Neurophysiology*, *122*(1), 43–51. https://doi.org/10.1016/j.clinph.2010.04.036

Gangi, D. N., Schwichtenberg, A. J., Ghent, S. B., Solis, E., Young, G. S., Iosif, A.-M., & Ozonoff, S. (2021). *Declining Gaze to Faces in Infants Developing Autism Spectrum Disorder Evidence From*. Child Development.

Gangi, D. N., Schwichtenberg, A. J., Iosif, A. M., Young, G. S., Baguio, F., & Ozonoff, S. (2018). Gaze to faces across interactive contexts in infants at heightened risk for autism. *Autism*, *22*(6), 763–768. https://doi.org/10.1177/1362361317704421

Gauthier, I., & Nelson, C. A. (2001). The development of face expertise. *Current Opinion in Neurobiology*, *11*(2), 219–224. https://doi.org/10.1016/S0959-4388(00)00200-2

Geangu, E., Quadrelli, E., Lewis, J. W., Macchi Cassia, V., & Turati, C. (2015). By the sound of it. An ERP investigation of human action sound processing in 7-month-old infants. *Developmental Cognitive Neuroscience*, *12*, 134–144. https://doi.org/10.1016/j.dcn.2015.01.005

Glauser, J., Wilkinson, C. L., Gabard-Durnam, L. J., Choi, B., Tager-Flusberg, H., & Nelson, C. A. (2022). Neural correlates of face processing associated with development of social communication in 12-month infants with familial risk of autism spectrum disorder. *Journal of Neurodevelopmental Disorders*, *14*(1), 1–14. https://doi.org/10.1186/s11689-021-09413-x

Gliga, T., Elsabbagh, M., Andravizou, A., & Johnson, M. (2009). Faces attract infants' attention in complex displays. *Infancy*, *14*(5), 550–562. https://doi.org/10.1080/15250000903144199

Goodwin, A., Hendry, A., Mason, L., Bazelmans, T., Begum Ali, J., Pasco, G., … Taylor, C. (2021). Behavioural measures of infant activity but not attention associate with later preschool adhd traits. *Brain Sciences*, *11*(5). https://doi.org/10.3390/brainsci11050524

Gredebäck, G., Astor, K., & Fawcett, C. (2018). Gaze following is not dependent on ostensive cues: A critical test of natural pedagogy. *Child development*, *89*(6), 2091-2098.

Gredebäck, G., Fikke, L., & Melinder, A. (2010). The development of joint visual attention: A longitudinal study of gaze following during interactions with mothers and strangers. *Developmental Science*, *13*(6), 839–848. https://doi.org/10.1111/j.1467-7687.2009.00945.x

Green, J., Nip, I., Wilson, E., Mefferd, A., & Yunusova, Y. (2010). Lip Movement Exaggerations During Infant-Directed Speech. *J Speech Lang Hear Res.*, *53*(6), 1529–1542. https://doi.org/10.1044/1092-4388(2010/09-0005

Greenstein, Y. J., Pavlides, C., & Winson, J. (1988). Lonf-term potentiation in the dentate gyrus is preferentially induced at theta rhythm periodicity. *Brain research*, *438*(1-2), 331-334.

Grossmann, T., & Johnson, M. H. (2007). The development of the social brain in human infancy. *The European Journal of Neuroscience*, *25*(4), 909–919. https://doi.org/10.1111/j.1460-9568.2007.05379.x

Grossmann, T., Johnson, M. H., Lloyd-Fox, S., Blasi, A., Deligianni, F., Elwell, C., & Csibra, G. (2008). Early cortical specialization for face-to-face communication in human infants. *Proceedings. Biological Sciences*, *275*(1653), 2803–2811. https://doi.org/10.1098/rspb.2008.0986

Grossmann, T., Johnson, M. H., Vaish, A., Hughes, D. A., Quinque, D., Stoneking, M., & Friederici, A. D. (2011). Genetic and neural dissociation of individual responses to emotional expressions in human infants. *Developmental Cognitive Neuroscience*, *1*(1), 57–66. https://doi.org/10.1016/j.dcn.2010.07.001

Gui, A., Bussu, G., Tye, C., Elsabbagh, M., Pasco, G., Charman, T., … Jones, E. J. H. (2021). Attentive brain states in infants with and without later Autism. *Translational Psychiatry*, *in press*. https://doi.org/10.1038/s41398-021-01315-9

Gui, A., Throm, E. V., da Costa, P. F., Haartsen, R., Leech, R., & Jones, E. J. H. (2022). Proving and improving the reliability of infant research with neuroadaptive Bayesian optimization. *Infant and Child Development*, (March), 1–6. https://doi.org/10.1002/icd.2323

Guo, J., Luo, X., Li, B., Chang, Q., Sun, L., & Song, Y. (2020). Abnormal modulation of theta oscillations in children with attention-deficit/hyperactivity disorder. *NeuroImage: Clinical*, *27*(June), 102314. https://doi.org/10.1016/j.nicl.2020.102314

Guy, M. W., Zieber, N., & Richards, J. E. (2016). The Cortical Development of Specialized Face Processing in Infancy. *Child Development*, *87*(5), 1581–1600. https://doi.org/10.1126/science.1249098.Sleep

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. https://doi.org/10.1016/j.dcn.2017.03.002

Haartsen, R., Charman, T., Pasco, G., Johnson, M. H., Jones, E. J. H., & Team, T. B. (2022). *Modulation of EEG theta by naturalistic social content is not altered in infants with family history of autism*. 1–14. https://doi.org/10.1038/s41598-022-24870-7

Hains, S. M. J., & Muir, D. W. (1996). Infant Sensitivity to Adult Eye Direction. *Child Development*, *67*(5), 1940–1951. https://doi.org/10.1016/S0163-6383(96)90043-0

Hakuno, Y., Hata, M., Naoi, N., Hoshino, E. ichi, & Minagawa, Y. (2020). Interactive live fNIRS reveals engagement of the temporoparietal junction in response to social contingency in infants. *NeuroImage*, *218*(August 2019), 116901. https://doi.org/10.1016/j.neuroimage.2020.116901

Halit, H., De Haan, M., & Johnson, M. H. (2003). Cortical specialisation for face processing: Face-sensitive event-related potential components in 3- and 12-month-old infants. *NeuroImage*, *19*(3), 1180–1193. https://doi.org/10.1016/S1053-8119(03)00076-4

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*(3), 300–308. https://doi.org/10.1080/17470919.2014.884982

Hoehl, S., & Striano, T. (2010). The development of emotional face and eye gaze processing. *Developmental Science*.

Hoehl, S., & Wahl, S. (2012). Recording infant ERP data for cognitive research. *Developmental Neuropsychology*, *37*(3), 187–209. https://doi.org/10.1080/87565641.2011.627958

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. https://doi.org/10.1016/j.dcn.2011.07.015

Ioannidis, J. P., Greenland, S., Hlatky, M. A., Khoury, M. J., Macleod, M. R., Moher, D., ... & Tibshirani, R. (2014). Increasing value and reducing waste in research design, conduct, and analysis. *The Lancet*, *383*(9912), 166-175.

Isaev, D. Y., Major, S., Murias, M., Carpenter, K. L. H., Carlson, D., Sapiro, G., & Dawson, G. (2020). Relative Average Look Duration and its Association with Neurophysiological Activity in Young Children with Autism Spectrum Disorder. *Scientific Reports*, *10*(1), 1912. https://doi.org/10.1038/s41598-020-57902-1

Jayaraman, S., Fausey, C. M., & Smith, L. B. (2015). The faces in infant-perspective scenes change over the first year of life. *PLoS ONE*, *10*(5), 13–15. https://doi.org/10.1371/journal.pone.0123780

Johnson, M. H. (2011). Interactive Specialization: A domain-general framework for human functional brain development? *Developmental Cognitive Neuroscience*, *1*(1), 7–21. https://doi.org/10.1016/j.dcn.2010.07.003

Johnson, M. H. (2017). Autism as an adaptive common variant pathway for human brain development. *Developmental Cognitive Neuroscience*, *25*, 5–11. https://doi.org/10.1016/j.dcn.2017.02.004

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, *40*(1–2), 1–19. https://doi.org/10.1016/0010-0277(91)90045-6

Johnson, M. H., Gliga, T., Jones, E., & Charman, T. (2015). Annual research review: Infant development, autism, and ADHD - Early pathways to emerging disorders. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *56*(3), 228–247. https://doi.org/10.1111/jcpp.12328

Johnson, M. H., Grossmann, T., & Kadosh, K. C. (2009). Mapping Functional Brain Development: Building a Social Brain Through Interactive Specialization. *Developmental Psychology*, *45*(1), 151–159. https://doi.org/10.1037/a0014548

Johnson, M. H., Jones, E. J. H., & Gliga, T. (2015). Brain adaptation and alternative developmental trajectories. *Development and Psychopathology*, *27*(2), 425–442. https://doi.org/10.1017/S0954579415000073

Johnson, M. H. (2001). *FUNCTIONAL BRAIN DEVELOPMENT IN HUMANS*. *2*(July).

Johnson, M. H, Griffin, R., Csibra, G., Halit, H., Farroni, T., De Haan, M., … Richards, J. (2005). The emergence of the social brain network: Evidence from typical and atypical development. *Development and Psychopathology*, *17*(3), 509–619. https://doi.org/10.1017/S0954579405050297

Jones, E. J.H., 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. https://doi.org/10.1038/s41598-020-67687-y

Jones, E. J.H. H, Venema, K., Earl, R., Lowy, R., Barnes, K., Estes, A., … Webb, S. J. (2016). Reduced engagement with social stimuli in 6-month-old infants with later autism spectrum disorder: A longitudinal prospective study of infants at high familial risk. *Journal of Neurodevelopmental Disorders*, *8*(1), 1–20. https://doi.org/10.1186/s11689-016-9139-8

Jones, E. J.H., Gliga, T., Bedford, R., Charman, T., & Johnson, M. H. (2014). Developmental pathways to autism: A review of prospective studies of infants at risk. *Neuroscience and Biobehavioral Reviews*, *39*, 1–33. https://doi.org/10.1016/j.neubiorev.2013.12.001

Jones, E. J.H. H, Dawson, G., Kelly, J., Estes, A., Jane Webb, S., Webb, S. J., & Jane Webb, S. (2017). Parent-delivered early intervention in infants at risk for ASD: Effects on electrophysiological and habituation measures of social attention. *Autism Research*, *10*(5), 961–972. https://doi.org/10.1002/aur.1754

Jones, E. J.H., Venema, K., Earl, R. K., Lowy, R., & Webb, S. J. (2017). Infant social attention: an endophenotype of ASD-related traits? *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *58*(3), 270–281. https://doi.org/10.1111/jcpp.12650

Jones, E. J.H., Venema, K., Lowy, R., Earl, R. K., & Webb, S. J. (2015). Developmental changes in infant brain activity during naturalistic social experiences. *Developmental Psychobiology*, *57*(7), 842–853. https://doi.org/10.1002/dev.21336

Jones, W., Carr, K., & Klin, A. (2008). Absence of preferential looking to the eyes of approaching adults predicts level of social disability in 2-year-old toddlers with autism spectrum disorder. *Archives of General Psychiatry*, *65*(8), 946–954. https://doi.org/10.1001/archpsyc.65.8.946

Jones, W., & Klin, A. (2013). Attention to eyes is present but in decline in 2-6-month-old infants later diagnosed with autism. *Nature*, *504*(7480), 427–431. https://doi.org/10.1038/nature12715

Karim, J., Weisz, R., & Rehman, S. U. (2011). International positive and negative affect schedule short-form (I-PANAS-SF): Testing for factorial invariance across cultures. *Procedia - Social and Behavioral Sciences*, *15*(2011), 2016–2022. https://doi.org/10.1016/j.sbspro.2011.04.046

Karras, T., Laine, S., & Aila, T. (2019). A style-based generator architecture for generative adversarial networks. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *2019-June*, 4396–4405. https://doi.org/10.1109/CVPR.2019.00453

Key, A. P. F., & Stone, W. L. (2012). Processing of novel and familiar faces in infants at average and high risk for autism. *Developmental Cognitive Neuroscience*, *2*(2), 244–255. https://doi.org/10.1016/j.dcn.2011.12.003

Klein, A., & Skrandies, W. (2013). A reliable statistical method to detect eyeblink-artefacts from electroencephalogram data only. *Brain Topography*, *26*(4), 558–568. https://doi.org/10.1007/s10548-013-0281-2

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a rKlimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. Brain Research Reviews, 29(2-3), 169–195. doi:10.1016/S016. *Brain Research Reviews*, *29*(2–3), 169–195. https://doi.org/10.1016/S0165-0173(98)00056-3

Klin, A., Shultz, S., & Jones, W. (2015). Social visual engagement in infants and toddlers with autism: Early developmental transitions and a model of pathogenesis. *Neuroscience and Biobehavioral Reviews*, *50*, 189–203. https://doi.org/10.1016/j.neubiorev.2014.10.006

Köster, M., Langeloh, M., Michel, C., & Hoehl, S. (2020). Young Infants Process Prediction Errors at the Theta Rhythm 1 2. *BioRxiv*, 2020.07.30.226902. https://doi.org/10.1101/2020.07.30.226902

Laeng, B., Sirois, S., & Gredebäck, G. (2012). Pupillometry: A window to the preconscious? *Perspectives on Psychological Science*, *7*(1), 18–27. https://doi.org/10.1177/1745691611427305

Lega, B. C., Jacobs, J., & Kahana, M. (2012). Human hippocampal theta oscillations and the formation of episodic memories. *Hippocampus*, *22*(4), 748–761. https://doi.org/10.1002/hipo.20937

Leong, V., Byrne, E., Clackson, K., Lam, S., & Wass, S. (2017). Speaker gaze changes 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. https://doi.org/10.1162/opmi_a_00008

Leppänen, J. M., Moulson, M. C., Vogel-Farley, V. K., & Nelson, C. A. (2007). An ERP Study of Emotional Face Processing in the Adult and Infant Brain. *Child Development*, *78*(1), 232–245. https://doi.org/10.1111/j.1467-8624.2007.00994.x

Leppänen, P. H. T., Guttorm, T. K., Pihko, E., Takkinen, S., Eklund, K. M., & Lyytinen, H. (2004). Maturational effects on newborn ERPs measured in the mismatch negativity paradigm. *Experimental Neurology*, *190*(SUPPL. 1), 91–101. https://doi.org/10.1016/j.expneurol.2004.06.002

Liu, H. M., Kuhl, P. K., & Tsao, F. M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science*, *6*(3), 1–10. https://doi.org/10.1111/1467-7687.00275

Lloyd-Fox, S., Blasi, A., Mercure, E., Elwell, C. E., & Johnson, M. H. (2012). The emergence of cerebral specialization for the human voice over the first months of life. *Social Neuroscience*, *7*(3), 317–330. https://doi.org/10.1080/17470919.2011.614696

Lloyd-Fox, S., Blasi, A., Pasco, G., Gliga, T., Jones, E. J. H., Murphy, D. G. M., … Yemane, F. (2018). Cortical responses before 6 months of life associate with later autism. *European Journal of Neuroscience*, *47*(6), 736–749. https://doi.org/10.1111/ejn.13757

Lloyd-Fox, S, Blasi, A., Elwell, C. E., Charman, T., Murphy, D., & Johnson, M. H. (2013). Reduced neural sensitivity to social stimuli in infants at risk for autism. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1758). https://doi.org/10.1098/rspb.2012.3026

Lloyd-Fox, Sarah, Széplaki-Köllod, B., Yin, J., & Csibra, G. (2015). Are you talking to me? Neural activations in 6-month-old infants in response to being addressed during natural interactions. *Cortex*, *70*, 35–48. https://doi.org/10.1016/j.cortex.2015.02.005

Lopera-Perez, D. C., Nayak, S., Stangier, G., Tuladhar, C. T., St. John, A. M., Grieve, P. G., & Tarullo, A. R. (2022). Infants' neural sensitivity to social interactions varies by income and infant-directed speech. *Developmental Psychobiology*, *64*(4), 1–14. https://doi.org/10.1002/dev.22265

Lord, C., Brugha, T. S., ..., & Veenstra-VanderWeele, J. (2022). Autism Spectrum Disorder. *Child and Adolescent Behavioral Health: A Resource for Advanced Practice Psychiatric and Primary Care Practitioners in Nursing*, *6*(1), 267–289. https://doi.org/10.1002/9781119487593.ch16

Lorenz, R., Hampshire, A., & Leech, R. (2017, March 1). Neuroadaptive Bayesian Optimization and Hypothesis Testing. *Trends in Cognitive Sciences*, Vol. 21, pp. 155–167. https://doi.org/10.1016/j.tics.2017.01.006

Lorenz, R., Johal, M., Dick, F., Hampshire, A., Leech, R., & Geranmayeh, F. (2021). A Bayesian optimization approach for rapidly mapping residual network function in stroke. *Brain*, *144*(7), 2120–2134. https://doi.org/10.1093/brain/awab109

Lorenz, R., Monti, R. P., Violante, I. R., Anagnostopoulos, C., Faisal, A. A., Montana, G., & Leech, R. (2016). The Automatic Neuroscientist: A framework for optimizing experimental design with closed-loop real-time fMRI. *NeuroImage*, *129*, 320–334. https://doi.org/10.1016/j.neuroimage.2016.01.032

Lorenz, R., Simmons, L. E., Monti, R. P., Arthur, J. L., Limal, S., Laakso, I., … Violante, I. R. (2019). Efficiently searching through large tACS parameter spaces using closed-loop Bayesian optimization. *Brain Stimulation*, *12*(6), 1484–1489. https://doi.org/10.1016/j.brs.2019.07.003

Lorenz, R., Violante, I. R., Monti, R. P., Montana, G., Hampshire, A., & Leech, R. (2018). Dissociating frontoparietal brain networks with neuroadaptive Bayesian optimization. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-

018-03657-3

Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT press.

Luyster, R., Gotham, K., Guthrie, W., Coffing, M., Petrak, R., Pierce, K., … Lord, C. (2009). The autism diagnostic observation schedule - Toddler module: A new module of a standardized diagnostic measure for autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *39*(9), 1305–1320. https://doi.org/10.1007/s10803-009-0746-z

Luyster, R. J., Powell, C., Tager-Flusberg, H., & Nelson, C. A. (2014). Neural measures of social attention across the first years of life: Characterizing typical development and markers of autism risk. *Developmental Cognitive Neuroscience*, *8*, 131–143. https://doi.org/10.1016/j.dcn.2013.09.006

Luyster, R. J., Wagner, J. B., Vogel-Farley, V., Tager-Flusberg, H., & Nelson, C. A. (2011). Neural correlates of familiar and unfamiliar face processing in infants at risk for autism spectrum disorders. *Brain Topography*, *24*(3–4), 220–228. https://doi.org/10.1007/s10548-011-0176-z

McClure, E., Chentsova-Dutton, Y., Holochwost, S., Parrott, W. G., & Barr, R. (2020). Infant emotional engagement in face-to-face and video chat interactions with their mothers. *Enfance*, *3*(3), 353–374. https://doi.org/10.3917/ENF2.203.0353

McDonald, N. M., & Perdue, K. L. (2018). The infant brain in the social world: moving toward interactive social neuroscience with functional near-infrared spectroscopy. *Neurosci Biobehav Rev*, *176*(87), 38–49. https://doi.org/10.1016/j.physbeh.2017.03.040

Menn, K. H., Michel, C., Meyer, L., Hoehl, S., & Männel, C. (2022). Natural infant-directed speech facilitates neural tracking of prosody. *NeuroImage*, *251*(September 2021), 118991. https://doi.org/10.1016/j.neuroimage.2022.118991

Meyer, M., Endedijk, H. M., van Ede, F., & Hunnius, S. (2019). Theta oscillations in 4-year-olds are sensitive to task engagement and task demands. *Scientific Reports*, *9*(1), 1–11. https://doi.org/10.1038/s41598-019-42615-x

Meyer, M., van Schaik, J. E., Poli, F., & Hunnius, S. (2022). How infant-directed actions enhance infants' attention, learning, and exploration: Evidence from EEG and computational modeling. *Developmental Science*, (March), 1–14. https://doi.org/10.1111/desc.13259

Miller, M., Iosif, A. M., Young, G. S., Hill, M. M., & Ozonoff, S. (2018). Early Detection of ADHD: Insights From Infant Siblings of Children With Autism. *Journal of Clinical Child and Adolescent Psychology*, *47*(5), 737–744. https://doi.org/10.1080/15374416.2016.1220314

Minagawa, Y., Xu, M., & Morimoto, S. (2018). Toward Interactive Social Neuroscience: Neuroimaging Real-World Interactions in Various Populations. *Japanese Psychological Research*, *60*(4), 196–224. https://doi.org/10.1111/jpr.12207

Moll, H., & Tomasello, M. (2004). 12- and 18-Month-Old Infants Follow Gaze To Spaces Behind Barriers. *Developmental Science*, *7*(1), 1–9. https://doi.org/10.1111/j.1467-7687.2004.00315.x

Moulson, M. C., Westerlund, A., Fox, N. A., Zeanah, C. H., & Nelson, C. A. (2009). The effects of early experience on face recognition: An event-related potential study of institutionalized children in romania. *Child Development*, *80*(4), 1039–

1056. https://doi.org/10.1111/j.1467-8624.2009.01315.x

Mullen, E. M. (1995). Mullen scales of early learning (AGS ed.). Circle Pines, MN: American Guidance Service Inc.

Mundy, P. (2018). A review of joint attention and social-cognitive brain systems in typical development and autism spectrum disorder. *European Journal of Neuroscience*, *47*(6), 497–514. https://doi.org/10.1111/ejn.13720

Mundy, P., Block, J., Delgado, C., Pomares, Y., Van Hecke, A. V., & Parlade, M. V. (2007). Individual differences and the development of joint attention in infancy. *Child Development*, *78*(3), 938–954. https://doi.org/10.1111/j.1467-8624.2007.01042.x

Mundy, P., Card, J., & Fox, N. (2000). EEG correlates of the development of infant joint attention skills. *Developmental Psychobiology*, *36*(4), 325–338. https://doi.org/10.1002/(SICI)1098-2302(200005)36:4<325::AID-DEV7>3.0.CO;2-F

Mundy, P., & Newell, L. (2009). Attention, Joint Attention, and Social Cognition. *Social Cognition*, *16*(January 2008), 269–274. https://doi.org/10.1111/j.1467-8721.2007.00518.x.Attention

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*(March 2017), 20–26. https://doi.org/10.1016/j.neuropsychologia.2017.03.030

Murias, M., Webb, S. J., Greenson, J., & Dawson, G. (2007). Resting State Cortical Connectivity Reflected in EEG Coherence in Individuals With Autism. *Biological Psychiatry*, *62*(3), 270–273. https://doi.org/10.1016/j.biopsych.2006.11.012

Murray, L., & Trevarthen, C. (1986). The infant's role in mother–infant communications. *Journal of child language*, *13*(1), 15-29.

Nelson, C. A., & de Haan, M. (1996). Neural correlates of infants' visual responsiveness to facial expression of emotion. *Developmental Psychobiology*, *29*(7), 577–595. https://doi.org/10.1002/(SICI)1098-2302(199611)29:7<577::AID-DEV3>3.0.CO;2-R

Nencheva, M. L., & Lew-Williams, C. (2022). Understanding why infant-directed speech supports learning: A dynamic attention perspective. *Developmental Review*, *66*(January), 101047. https://doi.org/10.1016/j.dr.2022.101047

Nguyen, T., Abney, D. H., Salamander, D., Bertenthal, B. I., & Hoehl, S. (2021). Proximity and touch are associated with neural but not physiological synchrony in naturalistic mother-infant interactions. *NeuroImage*, *244*(April), 118599. https://doi.org/10.1016/j.neuroimage.2021.118599

Nordt, M., Hoehl, S., & Weigelt, S. (2016). The use of repetition suppression paradigms in developmental cognitive neuroscience. *Cortex*, *80*, 61–75. https://doi.org/10.1016/j.cortex.2016.04.002

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(11), 2600–2606. https://doi.org/10.1073/pnas.1708274114

Nyström, P., Bölte, S., Falck-Ytter, T., Achermann, S., Andersson Konke, L., Brocki, K., … Zander, E. (2017). Responding to Other People's Direct Gaze: Alterations in Gaze Behavior in Infants at Risk for Autism Occur on Very Short Timescales. *Journal of Autism and Developmental Disorders*, *47*(11), 3498–

3509. https://doi.org/10.1007/s10803-017-3253-7

Nyström, P., Thorup, E., Bölte, S., & Falck-Ytter, T. (2019). Joint Attention in Infancy and the Emergence of Autism. *Biological Psychiatry*, *86*(8), 631–638. https://doi.org/10.1016/j.biopsych.2019.05.006

Okumura, Y., Kanakogi, Y., Kobayashi, T., & Itakura, S. (2020). Ostension affects infant learning more than attention. *Cognition*, *195*(September 2019), 104082. https://doi.org/10.1016/j.cognition.2019.104082

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. https://doi.org/10.1016/j.clinph.2005.12.027

Orekhova, Elena V., Stroganova, T. A., & Posikera, I. N. (1999). Theta synchronization during sustained anticipatory attention in infants over the second half of the first year of life. *International Journal of Psychophysiology*, *32*(2), 151–172. https://doi.org/10.1016/S0167-8760(99)00011-2

Orekhova, Elena V., Stroganova, T. A., & Posikera, I. N. (2001). Alpha activity as an index of cortical inhibition during sustained internally controlled attention in infants. *Clinical Neurophysiology*, *112*(5), 740–749. https://doi.org/10.1016/S1388-2457(01)00502-8

Ouyang, G., Dien, J., & Lorenz, R. (2022). Handling EEG artifacts and searching individually optimal experimental parameter in real time: A system development and demonstration. *Journal of Neural Engineering*, *19*(1). https://doi.org/10.1088/1741-2552/ac42b6

Ozonoff, S., Iosif, A.-M., Baguio, F., Cook, I. C., Hill, M. M., Hutman, T., … Young, G. S. (2010). A Prospective Study of the Emergence of Early Behavioral Signs of Autism. *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*(3), 256-266.e2. https://doi.org/10.1016/j.jaac.2009.11.009

Pellicano, E., & den Houting, J. (2022). Annual Research Review: Shifting from 'normal science' to neurodiversity in autism science. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *63*(4), 381–396. https://doi.org/10.1111/jcpp.13534

Petersen, S. E., & Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annual Review of Neuroscience*, *35*, 73–89. https://doi.org/10.1146/annurev-neuro-062111-150525

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. https://doi.org/10.1177/0956797619878698

Pierce, K., Conant, D., Hazin, R., Stoner, R., & Desmond, J. (2011). Preference for geometric patterns early in life as a risk factor for autism. *Archives of General Psychiatry*, *68*(1), 101–109. https://doi.org/10.1001/archgenpsychiatry.2010.113

Posner, M. I., & Petersen, S. E. (1990). The Attention System of the Human Brain. *Annals of Neuroscience*, 25.

Quaresima, V., & Ferrari, M. (2019). Functional Near-Infrared Spectroscopy (fNIRS) for Assessing Cerebral Cortex Function During Human Behavior in Natural/Social Situations: A Concise Review. *Organizational Research Methods*, *22*(1), 46–68. https://doi.org/10.1177/1094428116658959

Reeb-Sutherland, B. C., Levitt, P., & Fox, N. A. (2012). The predictive nature of individual differences in early associative learning and emerging social behavior. *PloS One*, *7*(1), e30511. https://doi.org/10.1371/journal.pone.0030511

Reynolds, G. D., Courage, M. L., & Richards, J. E. (2010). Infant Attention and Visual Preferences: Converging Evidence From Behavior, Event-Related Potentials, and Cortical Source Localization. *Developmental Psychology*, *46*(4), 886–904. https://doi.org/10.1037/a0019670.Infant

Reynolds, G. D., & Richards, J. E. (2005). Familiarization, Attention, and Recognition Memory in Infancy: An Event-Related Potential and Cortical Source Localization Study. *Dev. Psychol.*, *41*(4), 598–615. https://doi.org/10.3174/ajnr.A1256.Functional

Richards, J. E. (2003). Attention affects the recognition of briefly presented visual stimuli in infants: an ERP study. *Developmental Psychology*, *6*(3), 312–328. https://doi.org/10.3174/ajnr.A1256.Functional

Richards, J. E., Reynolds, G. D., & Courage, M. L. (2010). The neural bases of infant attention. *Current Directions in Psychological Science*, *19*(1), 41–46. https://doi.org/10.1177/0963721409360003.The

Richards, John E. (2001). Cortical Indexes of Saccade Planning Following Covert Orienting in 20-Week-Old Infants. *Infancy*, *2*(2), 135–157. https://doi.org/10.1207/S15327078IN0202_2

Richards, John E., Reynolds, G. D., & Courage, M. L. (2010). The neural bases of infant attention. *Current Directions in Psychological Science*, *19*(1), 41–46. https://doi.org/10.1177/0963721409360003

Rigato, S., Menon, E., Johnson, M. H., & Farroni, T. (2011). The interaction between gaze direction and facial expressions in newborns. *European Journal of Developmental Psychology*, *8*(5), 624–636. https://doi.org/10.1080/17405629.2011.602239

Rochat, P., Neisser, U., & Marian, V. (1998). Are young infants sensitive to interpersonal contingency? *Infant Behavior and Development*, *21*(2), 355–366. https://doi.org/10.1016/S0163-6383(98)90012-1

Rommelse, N. N. J., Franke, B., Geurts, H. M., Hartman, C. A., & Buitelaar, J. K. (2010). Shared heritability of attention-deficit/hyperactivity disorder and autism spectrum disorder. *European Child and Adolescent Psychiatry*, *19*(3), 281–295. https://doi.org/10.1007/s00787-010-0092-x

Rudling, M., Nyström, P., Bölte, S., & Falck-Ytter, T. (2022). Larger pupil dilation to nonsocial sounds in infants with subsequent autism diagnosis. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *63*(7), 793–801. https://doi.org/10.1111/jcpp.13520

Rutishauser, U., Ross, I. B., Mamelak, A. N., & Schuman, E. M. (2010). Human memory strength is predicted by theta-frequency phase-locking of single neurons. *Nature*, *464*(7290), 903–907. https://doi.org/10.1038/nature08860

Saby, J. N., & Marshall, P. J. (2012, April 1). The utility of EEG band power analysis in the study of infancy and early childhood. *Developmental Neuropsychology*, Vol. 37, pp. 253–273. https://doi.org/10.1080/87565641.2011.614663

Sainsbury, W. J., Carrasco, K., Whitehouse, A. J., McNeil, L., & Waddington, H. (2023). Age of diagnosis for co-occurring autism and attention deficit hyperactivity disorder during childhood and adolescence: A systematic review. *Review Journal of Autism and Developmental Disorders*, *10*(3), 563-575.

Salley, B., & Colombo, J. (2016). Conceptualizing Social Attention in

Developmental Research. *Social Development*, *25*(4), 687–703. https://doi.org/10.1111/sode.12174

Sandre, A., Freeman, C., Renault, H., Humphreys, K. L., & Weinberg, A. (2022). Maternal symptoms of depression and anxiety during the postpartum period moderate infants' neural response to emotional faces of their mother and of female strangers. *Cognitive, Affective and Behavioral Neuroscience*, *22*(6), 1370–1389. https://doi.org/10.3758/s13415-022-01022-y

Sauseng, P., Griesmayr, B., Freunberger, R., & Klimesch, W. (2010). Control mechanisms in working memory: A possible function of EEG theta oscillations. *Neuroscience and Biobehavioral Reviews*, *34*(7), 1015–1022. https://doi.org/10.1016/j.neubiorev.2009.12.006

Schwartz, S., Shinn-Cunningham, B., & Tager-Flusberg, H. (2018). Meta-analysis and systematic review of the literature characterizing auditory mismatch negativity in individuals with autism. *Neuroscience and Biobehavioral Reviews*, *87*(February), 106–117. https://doi.org/10.1016/j.neubiorev.2018.01.008

Schwier, C., van Maanen, C., Carpenter, M., & Tomasello, M. (2006). Rational imitation in 12-month-old infants. *Infancy*, *10*(3), 303–311. https://doi.org/10.1207/s15327078in1003_6

Senju, A., Csibra, G., & Johnson, M. H. (2008). Understanding the referential nature of looking: Infants' preference for object-directed gaze. *Cognition*, *108*(2), 303–319. https://doi.org/10.1016/j.cognition.2008.02.009

Senju, Atsushi, & Csibra, G. (2008). Gaze Following in Human Infants Depends on Communicative Signals. *Current Biology*, *18*(9), 668–671. https://doi.org/10.1016/j.cub.2008.03.059

Senju, Atsushi, Johnson, M. H., & Csibra, G. (2006). The development and neural basis of referential gaze perception. *Social Neuroscience*, *1*(3–4), 220–234. https://doi.org/10.1080/17470910600989797

Shephard, E., Bedford, R., Milosavljevic, B., Gliga, T., Jones, E.J.H., Pickles, A., Johnson, M.H., Charman, T. and (2019), Early developmental pathways to childhood symptoms of attention-deficit hyperactivity disorder, anxiety and autism spectrum disorder. J Child Psychol Psychiatr, 60: 963-974. https://doi.org/10.1111/jcpp.12947

Shephard, E., Fatori, D., Mauro, L. R., de Medeiros Filho, M. V., Hoexter, M. Q., Chiesa, A. M., … Polanczyk, G. V. (2019). Effects of Maternal Psychopathology and Education Level on Neurocognitive Development in Infants of Adolescent Mothers Living in Poverty in Brazil. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *4*(10), 925–934. https://doi.org/10.1016/j.bpsc.2019.05.009

Shimada, S., & Hiraki, K. (2006). Infant's brain responses to live and televised action. *NeuroImage*, *32*(2), 930–939. https://doi.org/10.1016/j.neuroimage.2006.03.044

Smith, E. S., Elliott, D., Killick, R., Crawford, T. J., Kidby, S., & Reid, V. M. (2021). Infants Oscillatory Frequencies change during Free-Play. *Infant Behavior and Development*, *64*(June 2020), 101612. https://doi.org/10.1016/j.infbeh.2021.101612

Stets, M., Stahl, D., & Reid, V. . (2012). Improving tools for the analysis of brain based measures of infant information processing abstract. *Developmental Neuropsychology*, *37*, 226–252.

Stormark, K. M., & Braarud, H. C. (2004). Infants' sensitivity to social contingency:

A “double video” study of face-to-face communication between 2- and 4-month-olds and their mothers. *Infant Behavior and Development*, *27*(2), 195–203. https://doi.org/10.1016/j.infbeh.2003.09.004

Striano, T. (2001). From social expectations to social cognition in early infancy. *Bulletin of the Menninger Clinic*, *65*(3), 361–370. https://doi.org/10.1521/bumc.65.3.361.19854

Striano, Tricia, Chen, X., Cleveland, A., & Bradshaw, S. (2006). Joint attention social cues influence infant learning. *European Journal of Developmental Psychology*, *3*(3), 289–299. https://doi.org/10.1080/17405620600879779

Striano, Tricia, Henning, A., & Stahl, D. (2006). Sensitivity to interpersonal timing at 3 and 6 months of age. *Interaction Studies. Social Behaviour and Communication in Biological and Artificial SystemsInteraction Studies / Social Behaviour and Communication in Biological and Artificial SystemsInteraction Studies*, *7*(2), 251–271. https://doi.org/10.1075/is.7.2.08str

Striano, Tricia, Reid, V. M., & Hoehl, S. (2006). Neural mechanisms of joint attention in infancy. *European Journal of Neuroscience*, *23*(10), 2819–2823. https://doi.org/10.1111/j.1460-9568.2006.04822.x

Stroganova, T. A., V. Orekhova, E., & Posikera, I. N. (1998). Externally and internally controlled attention in infants: An EEG study. *International Journal of Psychophysiology*, *30*(3), 339–351. https://doi.org/10.1016/S0167-8760(98)00026-9

Swingler, M. M., & Carver, L. J. (2013). Brain–Behavior Correlations: Relationships Between Mother– Stranger Face Processing and Infants’ Behavioral Responses to a Separation From Mother Margaret. *Developmental Psychology*, *46*(3), 669–680. https://doi.org/10.1037/a0018907.Brain

Swingler, M. M., Sweet, M. A., & Carver, L. J. (2007). Relations Between Mother-Child Interaction and the Neural Correlates of Face Processing in 6-Month-Olds. *Infancy*, *11*(1), 63–86. https://doi.org/10.1207/s15327078in1101_3

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

Tang, A., Lahat, A., Crowley, M. J., Wu, J., & Schmidt, L. A. (2019). Neurodevelopmental Differences to Social Exclusion: An Event-Related Neural Oscillation Study of Children, Adolescents, and Adults. *Emotion*, *19*(3), 520–532. https://doi.org/10.1037/emo0000456

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, *7*(1), 53–71. https://doi.org/10.1207/s15327078in0701_5

Tomasello, M. (1995). Language is not an instinct. *Cognitive Development*, *10*(1), 131–156. https://doi.org/10.1016/0885-2014(95)90021-7

Tomasello, M., & Carpenter, M. (2007). Shared intentionality. *Developmental Science*, *10*(1), 121–125. https://doi.org/10.1111/j.1467-7687.2007.00573.x

Trevarthen, C. (1985). Facial expressions of emotion in mother-infant interaction. *Human Neurobiology*, *4*(1), 21–32. Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-0021982567&partnerID=40&md5=cae5b2f48bacc6c4524db290e5f59d8a

Tummeltshammer, K., Feldman, E. C. H., & Amso, D. (2019). Using pupil dilation, eye-blink rate, and the value of mother to investigate reward learning mechanisms in infancy. *Developmental Cognitive Neuroscience*, *36*(May 2018), 100608. https://doi.org/10.1016/j.dcn.2018.12.006

Tye, C., Bussu, G., Gliga, T., Elsabbagh, M., Pasco, G., Johnsen, K., … Cura, M. (2020). *Understanding the nature of face processing in early autism: A prospective study Author Note*. https://doi.org/10.1101/2020.05.06.20092619

Tye, C., Rijsdijk, F., & McLoughlin, G. (2014). Genetic overlap between ADHD symptoms and EEG theta power. *Brain and Cognition*, *87*(1), 168–172. https://doi.org/10.1016/j.bandc.2014.03.010

Urakawa, S., Takamoto, K., Ishikawa, A., Ono, T., & Nishijo, H. (2015). Selective Medial Prefrontal Cortex Responses During Live Mutual Gaze Interactions in Human Infants: An fNIRS Study. *Brain Topography*, *28*(5), 691–701. https://doi.org/10.1007/s10548-014-0414-2

van der Molen, M. J. W., Dekkers, L. M. S., Westenberg, P. M., van der Veen, F. M., & van der Molen, M. W. (2017). Why don't you like me? Midfrontal theta power in response to unexpected peer rejection feedback. *NeuroImage*, *146*, 474–483. https://doi.org/10.1016/j.neuroimage.2016.08.045

van der Velde, B., & Junge, C. (2020). Limiting data loss in infant EEG: putting hunches to the test. *Developmental Cognitive Neuroscience*, *45*, 2020. https://doi.org/10.1016/j.dcn.2020.100809

van der Velde, B., White, T., & Kemner, C. (2021). The emergence of a theta social brain network during infancy. *NeuroImage*, *240*(May), 118298. https://doi.org/10.1016/j.neuroimage.2021.118298

van Noordt, S., Desjardins, J. A., Elsabbagh, M., Baron-Cohen, S., Bedford, R., Bolton, P., … Volein, A. (2022). Inter-trial theta phase consistency during face processing in infants is associated with later emerging autism. *Autism Research*, *15*(5), 834–846. https://doi.org/10.1002/aur.2701

van Noordt, S. J. R., White, L. O., Wu, J., Mayes, L. C., & Crowley, M. J. (2015). Social exclusion modulates event-related frontal theta and tracks ostracism distress in children. *NeuroImage*, *118*, 248–255. https://doi.org/10.1016/j.neuroimage.2015.05.085

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.

Viktorsson, C., Portugal, A. M., Li, D., Rudling, M., Siqueiros Sanchez, M., Tammimies, K., … Falck-Ytter, T. (2022). Preferential looking to eyes versus mouth in early infancy: heritability and link to concurrent and later development. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. https://doi.org/10.1111/jcpp.13724

Visser, J. C., Rommelse, N. N. J., Greven, C. U., & Buitelaar, J. K. (2016). Autism spectrum disorder and attention-deficit/hyperactivity disorder in early childhood: A review of unique and shared characteristics and developmental antecedents. *Neuroscience and Biobehavioral Reviews*, *65*, 229–263. https://doi.org/10.1016/j.neubiorev.2016.03.019

Vo, M. L. H., Smith, T. J., Mital, P. K., & Henderson, J. M. (2012). Do the eyes really have it? Dynamic allocation of attention when viewing moving faces. *Journal of Vision*, *12*(13), 1–14. https://doi.org/10.1167/12.13.3

Wan, M. W., Green, J., & Scott, J. (2019). A systematic review of parent–infant interaction in infants at risk of autism. *Autism*, *23*(4), 811–820. https://doi.org/10.1177/1362361318777484

Wass, S. V., Jones, E. J. H., Gliga, T., Smith, T. J., Charman, T., Johnson, M. H., … Volein, A. (2015). Shorter spontaneous fixation durations in infants with later emerging autism. *Scientific Reports*, *5*(Lc), 1–8. https://doi.org/10.1038/srep08284

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. https://doi.org/10.1371/journal.pbio.2006328

Webb, Sara J., Long, J. D., & Nelson, C. A. (2005). A longitudinal investigation of visual event-related potentials in the first year of life. *Developmental Science*, *8*(6), 605–616. https://doi.org/10.1111/j.1467-7687.2005.00452.x

Webb, Sara Jane, Jones, E. J. H., Merkle, K., Namkung, J., Toth, K., Greenson, J., … Dawson, G. (2010). Toddlers with elevated autism symptoms show slowed habituation to faces. *Child Neuropsychology*, *16*(3), 255–278. https://doi.org/10.1080/09297041003601454

Webb, Sara Jane, Jones, E. J. H., Merkle, K., Venema, K., Greenson, J., Murias, M., & Dawson, G. (2011). Developmental Change in the ERP Responses to Familiar Faces in Toddlers With Autism Spectrum Disorders Versus Typical Development. *Child Development*, *82*(6), 1868–1886. https://doi.org/10.1111/j.1467-8624.2011.01656.x

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: The SHINE toolbox. *Behavior Research Methods*, *42*(3), 671–684. https://doi.org/10.3758/BRM.42.3.671

Xie, W., McCormick, S. A., Westerlund, A., Bowman, L. C., & Nelson, C. A. (2019). Neural correlates of facial emotion processing in infancy. *Developmental Science*, *22*(3), e12758. https://doi.org/10.1111/desc.12758

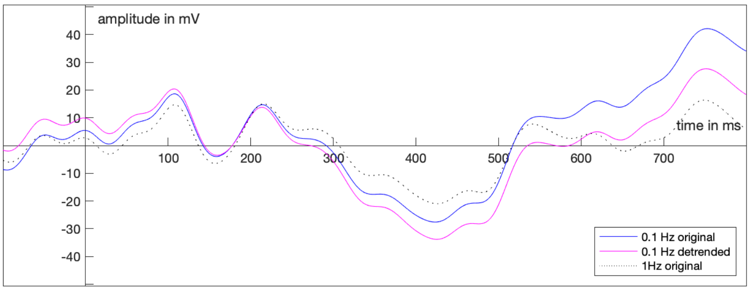
Yang, H. H., Savostyanov, A. N., Tsai, A. C., & Liou, M. (2011). Face recognition in Asperger syndrome: A study on EEG spectral power changes. *Neuroscience Letters*, *492*(2), 84–88. https://doi.org/10.1016/j.neulet.2011.01.061

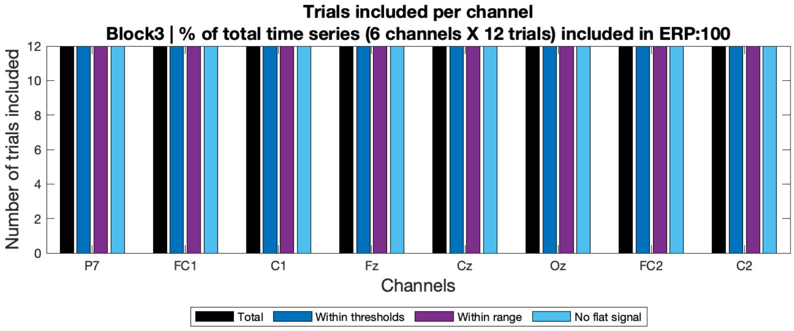
Zangl, R., & Mills, D. L. (2007). Increased brain activity to infant-directed speech in 6- and 13-month-old infants. *Infancy*, *11*(1), 31–62. https://doi.org/10.1207/s15327078in1101_2

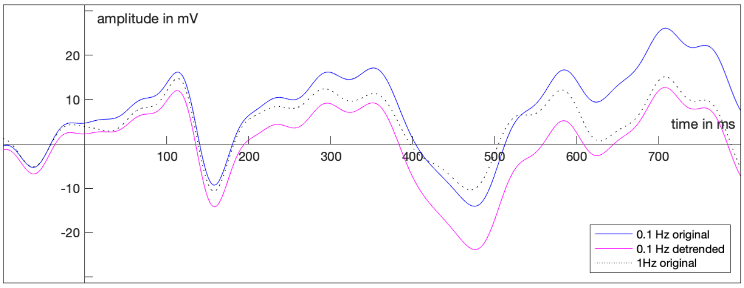
Zhang, Y., Koerner, T., Miller, S., Grice-Patil, Z., Svec, A., Akbari, D., … Carney, E. (2011). Neural coding of formant-exaggerated speech in the infant brain. *Developmental Science*, *14*(3), 566–581. https://doi.org/10.1111/j.1467-7687.2010.01004.x

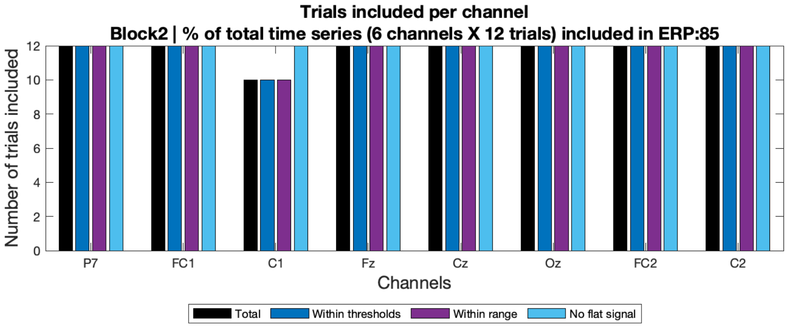
Appendix

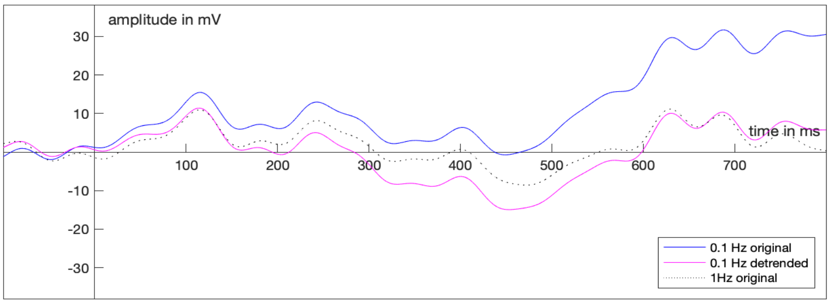
A.2.1. Comparisons of the shape of the individual block-level Nc when applying high pass filter of 1 Hz versus 0.1 Hz and detrending versus no detrending of the signal

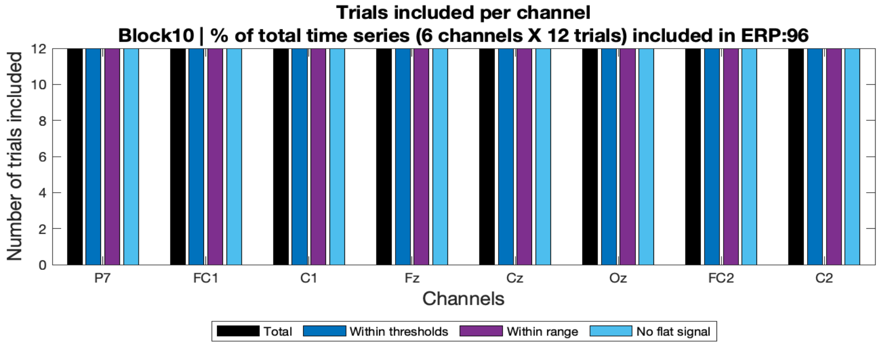


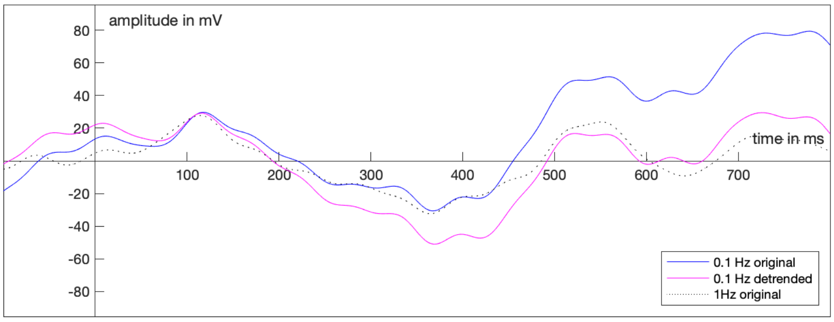


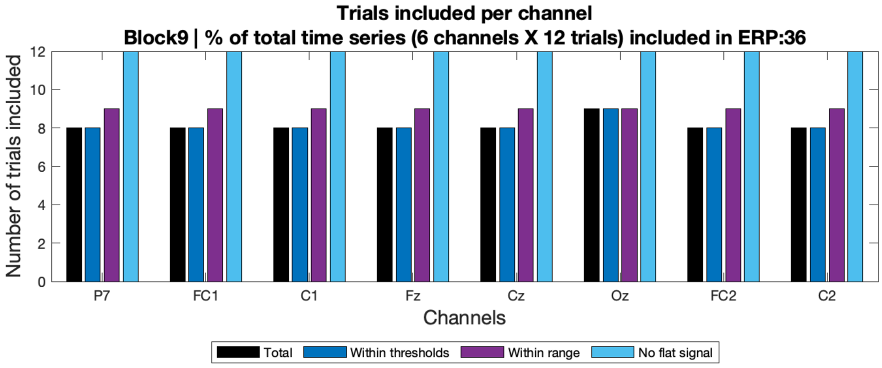


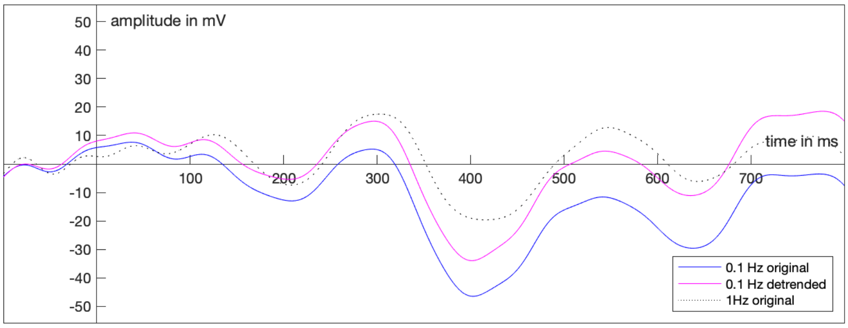


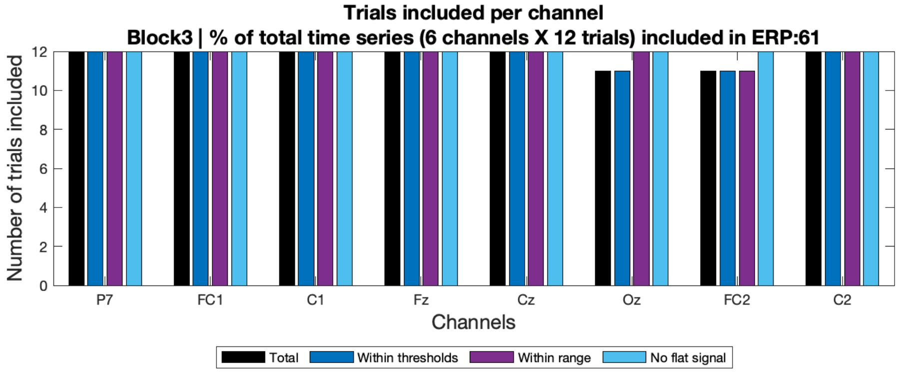

amplitude in mV
30
20
10
-10
-20
-30
100
200
300
400
500
600
700
time in ms
0.1 Hz original
0.1 Hz detrended
1Hz original

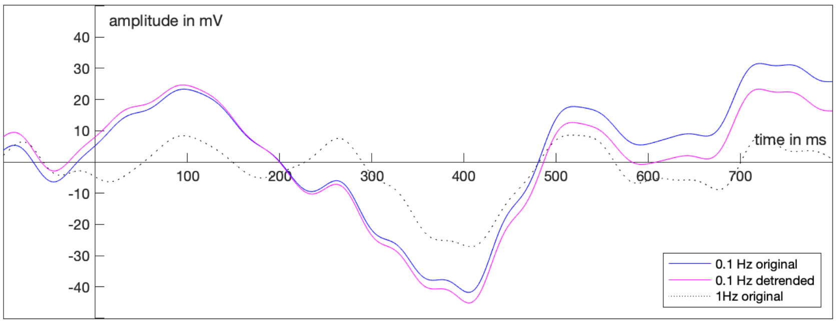

Trials included per channel
Block10 | % of total time series (6 channels X 12 trials) included in ERP:96
Number of trials included
12
10
8
6
4
2
0
P7
FC1
C1
Fz
Cz
Oz
FC2
C2
Channels
Total
Within thresholds
Within range
No flat signal

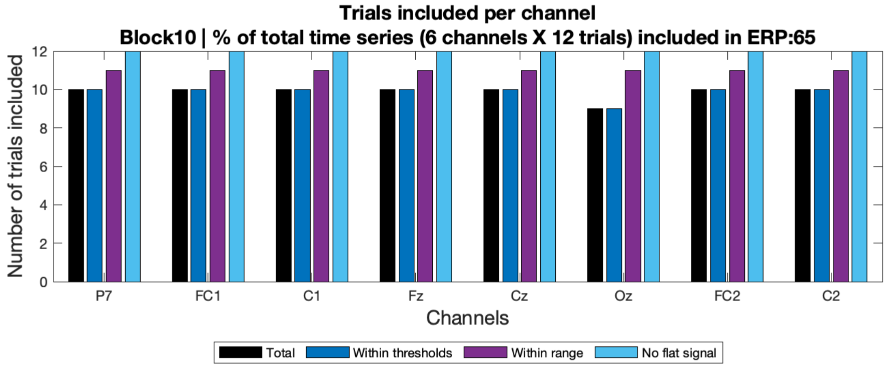

amplitude in mV
80
60
40
20
-20
-40
-60
-80
100
200
300
400
500
600
700
time in ms
0.1 Hz original
0.1 Hz detrended
1Hz original


Trials included per channel
Block9 | % of total time series (6 channels X 12 trials) included in ERP:36
Number of trials included
12
10
8
6
4
2
0
P7
FC1
C1
Fz
Cz
Oz
FC2
C2
Channels
Total
Within thresholds
Within range
No flat signal

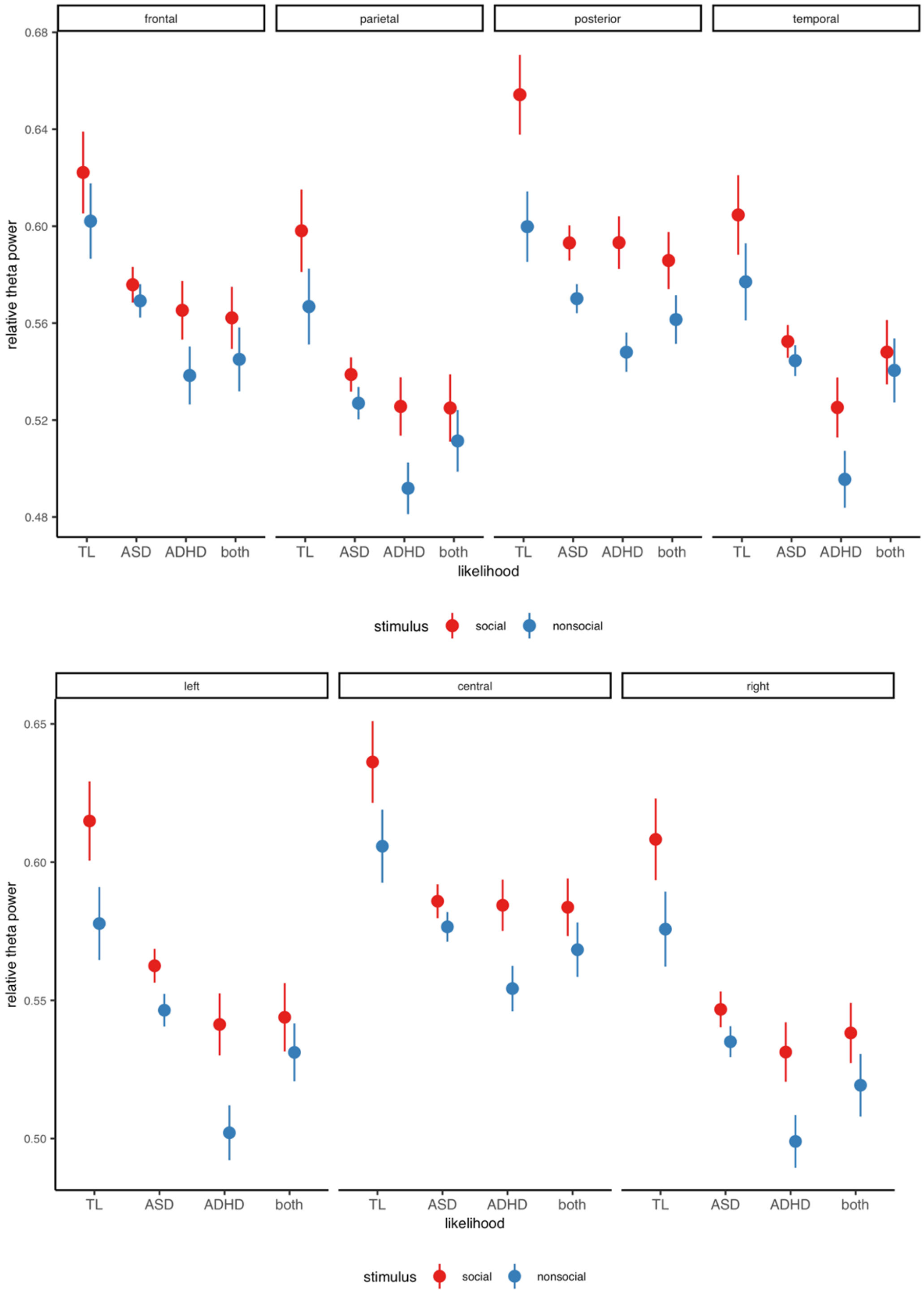



Trials included per channel
Block3 | % of total time series (6 channels X 12 trials) included in ERP:61

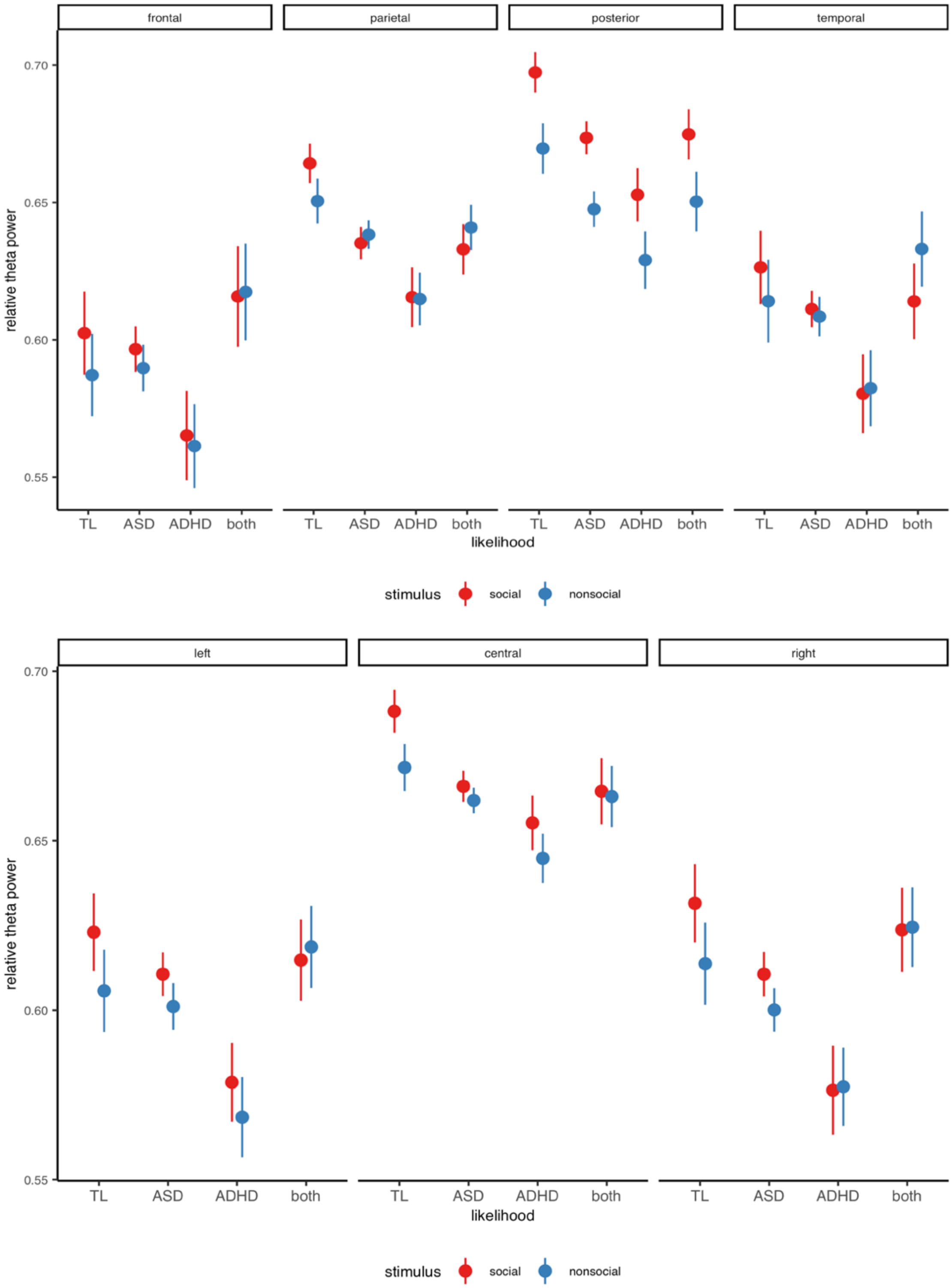



Trials included per channel
Block10 | % of total time series (6 channels X 12 trials) included in ERP:65

A.3.1. Modulation of social versus non-social theta power by likelihood status for different cortical regions (top) and hemisphere-side (bottom) at 14 months



A.4.1 Modulation of social versus non-social theta power by likelihood status for different cortical regions (top) and hemisphere-side (bottom) at 5 months

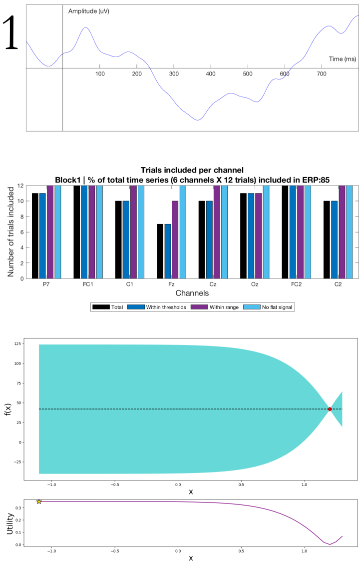


A.5.1 Individual output of an NBO experiment with a 1-dimensional stimulus space

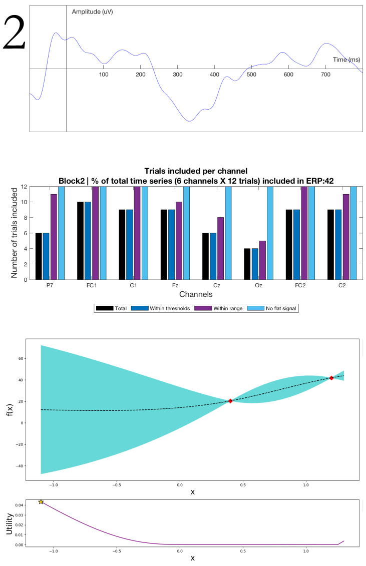
Graphical individual output of the NBO experiment (each panel representing one stimulus block). Top row: Plot of Nc waveform; middle row: feedback on the number of trials included per channel in the ERP, bottom row: objective function, with shaded area representing uncertainty, red dots representing empirical values, additional function representing acquisition function indexing the next stimulus point to sample

(see next page for figure)

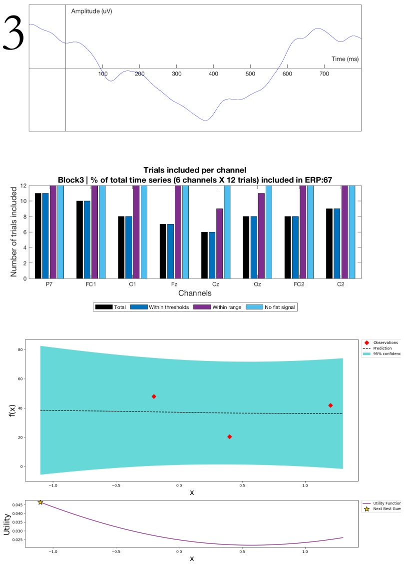
1


Amplitude (uV)
Time (ms)
100
200
300
400
500
600
700
Trials included per channel
Block1 | % of total time series (6 channels X 12 trials) included in ERP:85
Number of trials included
Channels
P7
FC1
C1
Fz
Cz
Oz
FC2
C2
Total
Within thresholds
Within range
No flat signal
f(x)
x
Utility
x

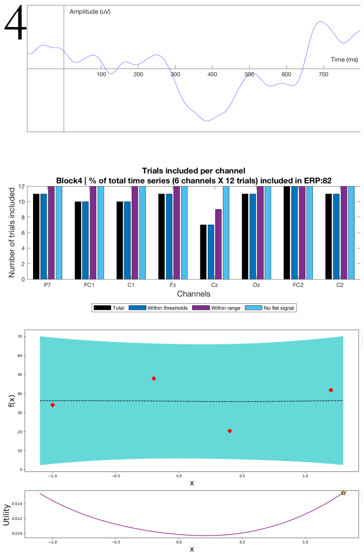
2


Amplitude (uV)
Time (ms)
100
200
300
400
500
600
700
Trials included per channel
Block2 | % of total time series (6 channels X 12 trials) included in ERP:42
Number of trials included
Channels
P7
FC1
C1
Fz
Cz
Oz
FC2
C2
Total
Within thresholds
Within range
No flat signal
f(x)
x
Utility
x

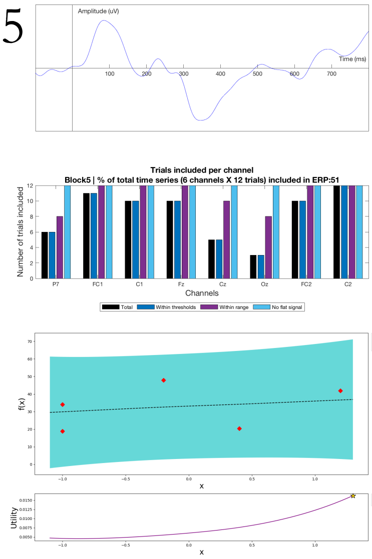
3


Amplitude (uV)
Time (ms)
100
200
300
400
500
600
700
Trials included per channel
Block3 | % of total time series (6 channels X 12 trials) included in ERP:67
Number of trials included
Channels
P7
FC1
C1
Fz
Cz
Oz
FC2
C2
Total
Within thresholds
Within range
No flat signal
Observations
Prediction
95% confidenc
f(x)
x
Utility Function
Next Best Gue
Utility
x

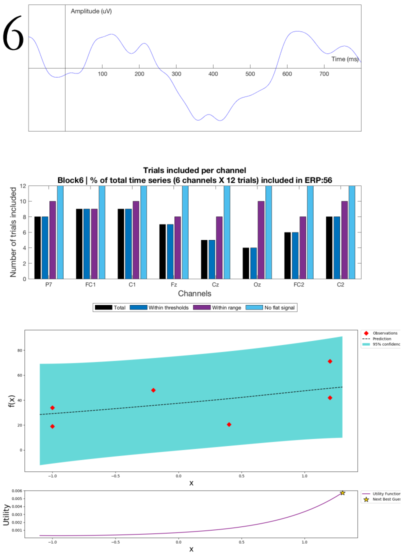
4


Amplitude (uV)
Time (ms)
100
200
300
400
500
600
700
Trials included per channel
Block4 | % of total time series (6 channels X 12 trials) included in ERP:82
Number of trials included
Channels
P7
FC1
C1
Fz
Cz
Oz
FC2
C2
Total
Within thresholds
Within range
No flat signal
f(x)
x
Utility
x

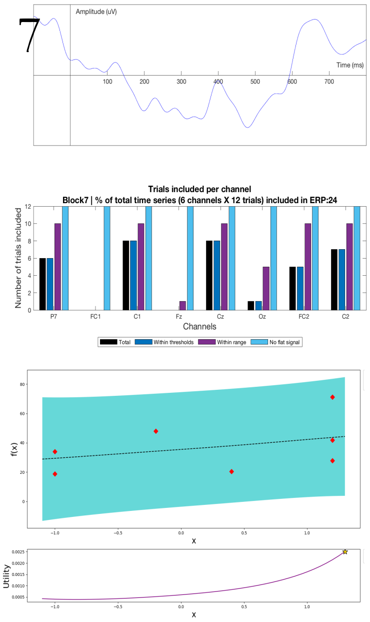
5


Amplitude (uV)
Time (ms)
100
200
300
400
500
600
700
Trials included per channel
Block5 | % of total time series (6 channels X 12 trials) included in ERP:51
Number of trials included
Channels
P7
FC1
C1
Fz
Cz
Oz
FC2
C2
Total
Within thresholds
Within range
No flat signal
f(x)
x
Utility
x

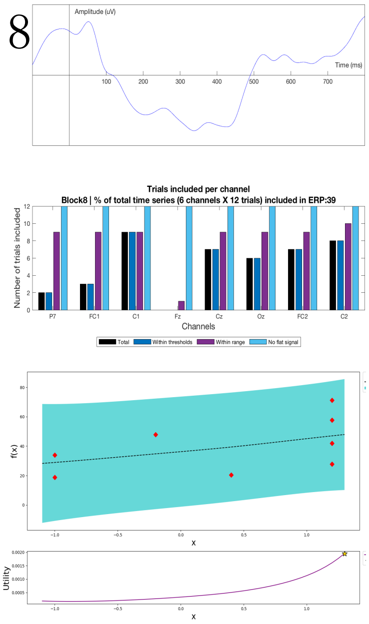
6


Amplitude (uV)
Time (ms)
100
200
300
400
500
600
700
Trials included per channel
Block6 | % of total time series (6 channels X 12 trials) included in ERP:56
Number of trials included
Channels
P7
FC1
C1
Fz
Cz
Oz
FC2
C2
Total
Within thresholds
Within range
No flat signal
Observations
Prediction
95% confidenc
f(x)
x
Utility Function
Next Best Gues
Utility
x

7


Amplitude (uV)
Time (ms)
100
200
300
400
500
600
700
Trials included per channel
Block7 | % of total time series (6 channels X 12 trials) included in ERP:24
Number of trials included
Channels
P7
FC1
C1
Fz
Cz
Oz
FC2
C2
Total
Within thresholds
Within range
No flat signal
f(x)
x
Utility
x

8


Amplitude (uV)
Time (ms)
100
200
300
400
500
600
700
Trials included per channel
Block8 | % of total time series (6 channels X 12 trials) included in ERP:39
Number of trials included
Channels
P7
FC1
C1
Fz
Cz
Oz
FC2
C2
Total
Within thresholds
Within range
No flat signal
f(x)
x
Utility
x

A.6.1 Live actor training procedure

The live actor training was delivered by the same researcher for all actors consisted of four steps:

1. (Remote or in-person) Actors were introduced to the stimulus space including the behaviours across the two dimensions and receiving overall instructions about body position during the experiment: *Your body should always be 45 degrees averted from child, while head and gaze angle are varied. Head and gaze angle are always varied together. During baseline, look at the screen (positioned at 90 deg to your right) and hold the toy still (i.e. no sound). Exhibit the same friendly facial expression throughout the experiment.*
2. (Remote or in-person) The trainer demonstrated each behaviour in the stimulus space.
3. (Remote or in-person) Actors practiced all behaviours.
4. (In-person) In mock-experiments in the lab, actors displayed the behaviours in response to instructions and received feedback from the trainer.

A.6.2 Theta power by VABS Interpersonal Relationships, by the four corners of the stimulus space

(NV = nonvocal; IDS = infant-directed speech)