Developing tools and expertise to include a diverse range of toddlers in developmental cognitive neuroscience research: mapping profiles of visual attention

Developing tools and expertise to include a diverse range of toddlers in developmental cognitive neuroscience research: Mapping profiles of visual attention

Eleanor K. Braithwaite

Thesis submitted for the degree of

Doctor of Philosophy

Department of Psychological Sciences

Birkbeck, University of London

October 2023

Acknowledgements

I’ve thought about what I’d write in my acknowledgements numerous times over the last few years, and it feels so rewarding to finally be writing them (though I seem to have lost the ability to write concisely!).

My thanks go first and foremost to my supervisor Emily Jones. Emily has been the most supportive and kind supervisor, and I couldn’t have hoped for anyone better under whose guidance to have started my research career. Emily has inspired me not just for her brilliant scientific contributions but also her approach to supervision of students and openness to look outside of the traditional view. The continuous opportunities and encouragement she directed my way have made such a huge difference to my research experiences and who I am as a researcher. Thank you for showing me that I might not only have a place but can also make important contributions to the research world.

To peers at Birkbeck and further afield, thank you. I’m grateful for all the conversations and interactions, however fleeting these may have been. A special thanks must go to the Safe Passage team in Cape Town, South Africa; my visit with you was certainly a highlight of my PhD. The openness and generosity I experienced made a huge impact on me personally and certainly helped to shape my work and this thesis. To my PhD peers, it’s been a weird few years and not the experience we hoped for, but I’m grateful for the experiences we have shared. To my lab group and particularly to Rianne and Teresa; thank you for your patience and kindness in answering my at-times endless questions and queries, and for showing me what it means to be a successful early career researcher.

A big thank you must go my friends, both those I’ve met in and outside of research. To my research pals Emily, Claire, and Britt – thank you for your support always but especially in the last few months, having you guys to discuss research & research life with has been such a source of reassurance. To Brittney especially, I am very glad that I sat next to you that very first masters lecture and made some excellent (!) small talk to break the ice. I’m so grateful I’ve had you one step ahead of me throughout this PhD, but more grateful to have you as a friend. Your work may have taken a different direction now, but I live in hope that we’ll one day co-author that paper we used to talk about together! To my friends outside of research who are too numerous to name, but to whom I owe a huge thanks. I am grateful for all the words of encouragement and especially

thankful for the time you have given listening to and supporting me even though the research world was so different from what you know. You guys are the best and I hope you know how grateful I am for each of you.

To my wonderful family who are genuinely the most supportive you could imagine, a million thankyous. To my amazing Mum – thank you for showing me what it means to do a job you love. It's always been clear to me the passion you have for your work and how important it is to you to help every single child that is lucky enough to come into your path. Your continual thriving to improve and be the best you can be is nothing short of amazing and I only hope you know how inspiring you are. Thanks also for always talking to me about your work, it's provided context and reminded me about the real-life impact of my research – and it's probably no coincidence that I study toddlers!

My lovely Dad - I know you won't believe me when I say it, but you also inspire me every day. A lot of people strive to be accepting and respectful of all, but there's very few who are as genuinely kind and open-minded as you. You have shown me what it means to be genuinely interested in everyone for who they are, regardless of where they're from or what they do. I am so appreciative of your undoubting support, belief, and interest in whatever I choose to pursue – and for the healthy dose of mocking that usually comes with it!

My big sister Hollie - never before and never again will there be a more supportive sibling in the world. You are my fiercest protector and my biggest cheerleader, and I am so thankful to have you to lean on. I'm so grateful we've been in London together throughout the last few years; you've been my piece of home away from home. You're my number one since day one and the reassurance I get from knowing you're always supporting even the most minor achievement is immense; thank you for always celebrating the little things with me, it means the world.

To my Tber – right from the start, you made me believe a PhD might be something I could do, and you have been my strength and reassurance in all my moments of doubt. You've been by my side in all the small and big moments, and I'm not sure I would made have it through without you - there will never be enough thanks for all that you've done for me. Thanks also for being willing to model our holidays around a conference or research placement - I know my PhD hasn't always been fun but I like to think we've had some

pretty good moments! I'm sorry I'm not pursuing the banking career you dreamed of for me (!) but I'm excited for what's come and so grateful to have you by my side for it all.

A final thank you to all the families who took part in my research and to friends and family who helped proofread parts of this thesis - it helped enormously with my panic in the final stages. Special thanks also to Hol for checking multiple chapters as well as cooking and delivering meals to me when Christopher was away!

Finally, I am very grateful for having had the opportunity to do this PhD, and I sincerely hope that the future sees individuals from all backgrounds being afforded a similar privilege.

Abstract

Given rapid brain and cognitive development during the first years of life, it is likely that experiences during this period could play a particularly significant role in development. Though research indicates a relation between socioeconomic status (SES) – relating to an individual's social standing - and both cognitive and brain development, to date much of this research has included samples that are biased toward including children from high socioeconomic families. This limits the generalisability of findings and might cause aspects of environment-development relations to be missed. Developmental Cognitive Neuroscience (DCN) additionally lacks specialism to conduct research with toddlers, meaning an understanding of neurocognitive development during this period is somewhat deficient. Such that relations between experiences and development during the early years of life can be fully understood, this thesis focussed on developing tools and expertise to fill these two gaps and to build knowledge of toddler neurodevelopment. Chapter 2 investigated how SES was related to profiles of visual attention in young toddlers. A data-driven approach to investigating SES measures found two clusters which largely related to typically low and high SES groups. Given that a battery of eye-tracking tasks was found to be a feasible method for collecting visual attention information from a large sample of 18-month-olds, eye-tracking data was used to compare performance between two SES groups. This revealed that the higher SES group looked less to faces, were faster to find a hidden object in a working memory paradigm and were faster and more accurate in the learning phase of a cognitive control task than the low SES group. The study in chapter 3 utilised a two-visit design to assess the reliability of a wearable, wireless electroencephalography (EEG) system with typically developing toddlers. It focussed specifically on alpha and theta power measures during video viewing, as these are thought to be neural measures involved with cognition and learning and may be particularly involved in aspects of development which are impacted by early environmental experiences. Numerous measures of theta and alpha power were calculated, with significant differences in EEG power found between different frequency bands, brain regions, video conditions and other comparisons. Reliability for these measures varied considerably, with relative theta power over the whole head and whole video providing the most reliable measures. Chapter 4 used a paradigm designed specifically to gather good quantity and quality EEG data from toddlers, which was additionally suited to use in less-controlled settings. It found that 2- and 3-year-olds' theta

power was higher during social and exploration, compared to non-social and bubble blowing conditions, whilst alpha was lower during exploration compared to all other conditions. Relations were also found between depth (length) of exploration and each of theta and alpha in posterior regions, whilst a positive relation between experience of chaotic environments and depth of exploration may be considered in context of useful adaptations to experiences. Feasibility analyses showed high parental acceptability and suitability of this less-controlled design for gathering neurocognitive data from toddlers. Chapter 5 investigated methods for increasing the diversity of participants in developmental research, through the development of a scalable app-based measure of early development. Focus groups and questionnaire data identified factors which parents consider important for research utilising an app-based tool, which could have influential implications for future development of this research. A current app-based tool revealed strong relations between app measures, age, and other cognitive measures, supporting the validity of this tool for cognitive data collection. Finally, a data-driven approach found two clusters among numerous SES variables which mapped to a low and high SES group as is typically used in research, though analyses did not reveal a relation between SES grouping and cognitive ability. Overall, this thesis has made both theoretical and practical contributions to knowledge of neurocognitive development in toddlers and has provided expertise for improving diversity in future DCN research.

Table of contents

Contents

List of figures

List of tables

List of abbreviations

ANOVA	Analysis of variance
AOI	Area of interest
AIC	Akaike information criterion
ASD	Autism Spectrum Disorder
ASQ-3	Ages and Stages Questionnaire – version 3
BIC	Bayesian information criterion
BIS	Behavioural inhibition system questionnaire
CBCL	Child Behaviour Checklist
CHAOS	Confusion hubbub and order scales
CIs	Confidence intervals
CMS	Common mode sense
CS	Central stimulus
CSPS	Cognitively stimulating parenting scale
DCN	Developmental cognitive neuroscience
DRL	Driven right leg
EBS	Exploratory Behaviour Scale

EEG	Electroencephalography
EF	Executive function
ERP	Event-related potential
EYFS	Early Years Foundation Stage
fMRI	Functional magnetic resonance imaging
fNIRS	Functional near-infrared spectoscropy
GAD-7	General Anxiety Disorder questionnaire
IBQ-R	Infant behavioural questionnaire revised
IMD	Index of multiple deprivation
ICC	Intraclass correlation coefficient
LB	Lower bound
LM	Linear model
LMM	Linear mixed model
MANOVA	Multivariate analysis of variance
MRI	Magnetic resonance imaging
PS	Peripheral stimulus
RMSEA	Root mean square error of approximation

RT	Reaction time
SEM	Structural equation model
SES	Socioeconomic status
SRT	Saccadic reaction time
UB	Upper bound

1 GENERAL INTRODUCTION

1.1 OPENING DISCUSSION

Every human being is an individual who experiences the World in a unique way. Though we may share aspects of experiences with others, the way they impact us is exclusively ours. Even identical twins who share parents and genetics do not experience the World in an identical way. Whilst many of these differences are subtle and intangible, more significant differences in experiences can be measured and quantified. A question which has been considered by many is how these experiences may impact an individual's development. Given the rapid changes that occur during children's early development, it is likely that their experiences during this time are particularly impactful.

This concept is not new to the field of Developmental Cognitive Neuroscience (DCN); indeed many campaigns such as Shaping Us (https://shapingus.centreforearlychildhood.org/; Centre for Early Childhood, 2023), The Early Years Foundation Stage Framework (Department for Education, 2021) and reference to the 'First 1000 days' (Health and Social Care Committee, 2019) demonstrate how this important period of development is being brought to public attention. The research work behind these campaigns, whilst contributing significantly to our understanding of how experience impacts development, is typically limited in two main ways. Both limitations are based around *who* is usually included in this research, meaning the work has limited generalisability across populations. Firstly, research studies typically include children from a limited range of backgrounds, with a bias for those from high socioeconomic families. Second is the lack of specialism in conducting research with toddlers, meaning neurodevelopmental research is not typically conducted with children of this age, instead favouring either younger infants or older children (i.e. preschoolers).

This thesis focusses on overcoming these two deficits, by developing tools and expertise for improving the diversity of participant samples and for including a greater number of toddlers in DCN research. It additionally explores the relation between early experiences and cognition in toddlerhood, making empirical contributions detection of experience-related differences in visual attention. It employs various methods and utilises technical advances to develop and assess study designs and approaches in relation to both these goals. This thesis therefore has both a theoretical and somewhat practical focus and aims to develop expertise and knowledge which will facilitate improved future research investigating the impact of experiences on early development.

This chapter will provide a background for the following chapters of this thesis, by first summarising the existing field of research before critically reviewing its contributions. This introduction will begin by considering development in the first years of life and highlighting why research which focusses on toddlers and increasing diversity is needed. Attention will then be introduced as a domain which may be particularly related to early environmental experiences, and work which looks at the relation between these will be discussed. The following section will include justification for using a neurocognitive approach to understand relations between early experiences and development, as well as a review of existing work which does this, with a particularly focus on attention-relevant neural measures. Section 1.5 includes some discussion of theories about how socioeconomic status (SES), attention and neural function may be connected, before a section which outlines existing gaps and limitations in existing literature. This includes challenges associated with defining and measuring early experiences – specifically focussing on SES - and choosing a theoretical framework of this. Finally, an overview of each ensuing chapter will be provided.

1.2 EARLY EXPERIENCES AND DEVELOPMENT

A considerable amount of work has explored the impact of early experiences on child development. In much of this work, early experiences have been encapsulated as socioeconomic status (SES). In general terms, SES can be described as a measure of an individual's standing within society which is typically described in relative terms as low, medium, or high. Though there is no consensus about exactly how SES should be conceptualised and measured (Bradley & Corwyn, 2002), it is usually determined by social and economic measures such as education, income and occupation. This section will outline what the existing expansive literature states about the relation between SES and development, with a particular focus on toddlerhood.

SES is commonly reported as one of the most significant predictors of functioning through life. There is evidence to support a link between SES and measures of health, socioeconomic and cognitive outcomes from even as early as prenatally to adulthood (Bradley & Corwyn, 2002; Conroy et al., 2010; Walpole, 2003). Within development, there is an abundance of work documenting the impact of SES, with evidence indicating links between lower SES and poorer language, memory, executive function (EF) and other cognitive abilities (Duncan et al., 1998; Hackman & Farah, 2009; Ursache & Noble, 2016). In fact, this effect is so apparent that the difference in abilities between children from low versus high SES backgrounds is commonly referred to as the school attainment gap and is a focus of UK education policy (*Education in England*, 2018). This gap is clearly discernible at school start age (Janus & Duku, 2007), with early differences persisting throughout school and beyond such that better cognitive, social and emotional abilities at school start tend to be associated with higher performance in later school measures (Duncan et al., 2007; Hair et al., 2006a; Li-Grining et al., 2010; McClelland et al., 2006; McWayne et al., 2004).

Children's abilities when they start school are commonly referred to as 'school readiness', which is a measure of how prepared an individual is for the school environment (Snow, 2006). Measures of school readiness typically include assessment of social and emotional, as well as cognitive abilities (Janus & Duku, 2007), whilst some work has also considered characteristics of the contexts and systems around a child (High & and the Committee on Early Childhood, 2008; Mashburn & Pianta, 2006). Regardless of how it is measured, there is evidence that better school readiness is associated with higher performance in later school measures (Hair et al., 2006b; Li-Grining et al., 2010; McClelland et al., 2006; Snow, 2006). Given these strong relations between children's abilities at school start and later skills, there is argument that supporting children's development and preparation before they start school could have substantial impacts on later abilities.

In considering the development of school readiness, perhaps the most significant factor is children's early environment (High & and the Committee on Early Childhood, 2008). A considerable body of research has found that children from more disadvantaged backgrounds may be less prepared to start school, with evidence of a link between different measures of both disadvantage and school readiness. For instance, Hair et al. (2006) found that a composite measure of household SES (incorporating income and parental education measures) predicted children's school readiness profiles, as did family status. Children from more disadvantaged backgrounds were more likely to have a profile of school readiness associated with lower language and cognition skills and higher risk of later health difficulties. When categorised by their family's poverty status (determined by measures of income and family size), kindergartners from families living below the poverty threshold have lower reading and maths scores than did peers who were above the poverty threshold (Denton et al., 2002). There is additionally evidence that worse nutrition (associated with lower SES) between 18-months- to 5-years-old may be related to lower school readiness across multiple domains including language, social and emotion (Omand et al., 2021).

Other work has indicated relations between SES and school readiness which may be mediated by other measures. Executive functioning (EF) – abilities that underpin cognitive processes crucial for attention, memory, planning and balancing multiple tasks - in particular, might be involved in the association between SES and school readiness. For instance, children from higher SES families (measured by a composite of education and occupation variables) were found to be better able to switch between tasks in changing environmental demands (an aspect of EF), with this higher set-shifting skill also associated with more advanced school readiness (Micalizzi et al., 2019). As well as indirect links, this work found a direct positive association between SES and school readiness and indicated that greater household chaos (associated with lower SES) was related to lower readiness for school. Other work also found that preschoolers' EF abilities partially accounted for the relation between SES (determined by enrolment on need-based or privately funded preschool education) and mathematic ability at school start. (Fitzpatrick et al., 2014). In addition to EF, some research has indicated that parenting behaviours might be involved in the association between SES and school readiness. Specifically, there is evidence that aspects of parenting behaviours which mediate this relation may differ across racial/ ethnic groups (Dotterer et al., 2012), further impacting the significance that culture and experiences may have on young children even before school start age. Indeed, environmental experiences in the first years of life can have long-term impacts that last longer than just throughout schooling, and which affect health, behaviour and productivity throughout life (Shonkoff et al., 2012).

Whilst there is clearly a relationship between SES and development, understanding the details of this association is much more complex. In more recent years the focus has shifted from questions about *whether* SES is related to development to questions about *how* they are related. In this work, different levels of observation have been explored, with a combination of behavioural, cognitive and neuroimaging methods increasingly being used to explore the nature of this association. This DCN approach may facilitate insights which would not otherwise be apparent when using methods in isolation; insights which may ultimately enable challenges caused by the impact of SES to be ameliorated.

1.2.1 Why experiences are particularly important in the early years of life

Given that the school readiness gap is already apparent in children aged 3- and 4-years-olds ('preschoolers') (Passaretta et al., 2020; Passaretta & Skopek, 2018), it seems likely that a significant association between SES and development also exists in toddlerhood.

Indeed, consideration of neurobiological work further indicates that early brain development may be particularly impacted by environmental factors. This might occur not only through direct impacts, but experiences could also affect development via epigenetic pathways (Gilmore et al., 2018). For instance, there is evidence to suggest that gene expression can be impacted by environmental factors, with patterns and timing of gene expression then impacting early brain architecture (Fox et al., 2010). In addition to gene expression, some biological characteristics of brain development make it particularly able to change in response to early experiences. During the first few years of life, billions of new synapses are created at a rapid rate. Following this initial abundance of connections, synaptic pruning occurs such that lesser-used connections are removed, and neural circuitry becomes more efficient. Though there is some prescribed order to when broad brain regions undergo this process (Krishnan & Johnson, 2014; Tierney & Nelson, 2009), specific pruning is determined by early environments, with occurrences of particular experiences causing some neural connections to strengthen whilst others are lost (Council et al., 2000; Tierney & Nelson, 2009). This ability of the brain to alter and change in response to experiences is referred to as neural plasticity, with a rapid increase in synapses in the first years of life indicating this as a period of heightened plasticity (Tierney & Nelson, 2009).

Linked to neural plasticity is the concept of sensitive periods, which indicate the importance of timing in relations between early experiences and brain development. Indeed, sensitive periods refer to times when the brain is maximally sensitive to environments and are characterised by high neural plasticity (Tierney & Nelson, 2009). Sensitive periods for neural processing systems occur at different developmental points, with sensitive periods for basic visual and hearing systems occurring at a younger age than for higher cognitive functions (Johnson, 2001; Krishnan & Johnson, 2014). Though brain changes can occur outside of sensitive periods, this is harder, as the development of more specific and complex neural systems are established based on early neural tuning (Krishnan & Johnson, 2014). In this way, very early brain development is thought to have cascading effects for later brain and cognitive functioning, meaning that early experiences which shape this early development can have lasting consequences. In addition, these characteristics of neurobiological development indicate the early years of life as a period when environmental experiences may play a particularly important role in development.

Although the vast majority of work in this area has focussed on children from preschool age upwards, there is some evidence suggesting that the effects of SES begin during infancy (Fransson et al., 2007) and continue throughout toddlerhood (Fernald et al., 2013; von Stumm & Plomin, 2015).

1.3 ATTENTIONAL PROCESSES

One aspect of development that may be particularly fruitful to investigate when considering potential relations with SES is attention. Attentional processes are critical to how children perceive and understand the world, and are a key mechanism via which they acquire knowledge, with initially interests in infant attention perhaps sparked by findings that attention was related to other measures of cognitive ability (Stevens & Bavelier, 2012). In addition, attention abilities undergo significant development during the first years of life, with a shift from reliance upon exogenous-driven and non-specific orienting systems to more complex executive attention processes (Conejero & Rueda, 2020). This section discusses theories of attention, before presenting some of the existing research investigating the link between SES and attentional processes in early development.

1.3.1 Theories of attention

Posner & Petersen (1990) proposed a framework which included three networks of attention relating to orienting, alerting and executive attention. Alerting refers to the maintenance of vigilance, orienting is the ability to prioritise sensory information from the environment, whilst executive attention is related to control and management of conflict (Petersen & Posner, 2012). This attention network was influential when introduced for three main reasons; it suggested attention is not a unitary concept but a set of networks, that these networks each have individual functions, and that attention is its own concept which interacts with other systems (de Souza Almeida et al., 2021). Such ideas are no longer novel, with various work indicating multiple aspects of attention which may be underpinned by different neural networks exist (Chica et al., 2013; Posner et al., 2014). Indeed, numerous other researchers have considered environmentally-driven (exogenous) attention as distinct from internally-driven (endogenous) attention (Chica et al., 2013; Sarter et al., 2001), not dissimilar to the way Posner & Petersen (1990) separated executive attention from stimulus-driven orienting. It has been proposed that executive attention relates closely to endogenous attention, with a shared focus on resolving conflict and performing goal-related behaviours (Amso & Scerif, 2015; Miller & Cohen, 2001). Exogenous attention, on the other hand, may be driven by properties of a stimulus such as luminance or complexity.

Perhaps more recently, thoughts have turned towards how social attention may integrate with other aspects of attention. A considerable body of work discusses the term ‘social attention’, generally referring to attention related to social stimuli, yet there is little consensus on the theory behind this (see Braithwaite et al., 2020). Social attention may be a result of social motivation interacting with other aspects of attention, such as orienting or maintenance (Chevallier et al., 2012). Alternatively, social attention may be a special case of attention to object features which are socially-relevant. Much of the work considering this element of attention comes from research investigating where social attention may be atypical, such as in neurodiverse individuals with conditions such as autism spectrum disorder (ASD). Models of attention have not yet delineated the mechanisms of social attention, but various developmental work does indicate that it may be distinct from ‘non-social’ attention. Numerous studies have found differences in neural signals relating to social and non-social attention (Hoehl et al., 2014; Jones et al., 2015), thereby indicating some degree of separation in their function and neural basis. Tasks which assess social attention include stimulus relating to other people (i.e. faces, voices) whilst non-social tasks would exclude this as far as possible, instead using stimuli such as vehicles or toys.

In terms of development, there is evidence of considerable development of attentional processes in the first few years of life (Conejero & Rueda, 2020). Early in the first year, infants learn to sustain their attention to increasingly complex stimuli and to disengage from a focal stimulus, whilst executive attention systems in particular undergo rapid development during the second year of life (Conejero & Rueda, 2020). Behavioural observations of this development are coupled with work indicating changes at the neural level, with consolidation of brain areas involve in the executive attention system related to cognitive ability in toddlers (Alcauter et al., 2014). Evidence of executive skills, in the form of inhibitory control, seem to emerge from 9 months old (Holmboe et al., 2008), though it is only from 18 months old that this undergoes considerable improvement (see Conejero & Rueda, 2020 for overview).

A range of tasks have been used to assess different skills relating to attentional processing. The attention network task (ANT) was developed as a way to assess the alerting, orienting and executive control systems involved in attention (Fan et al., 2005), and has been used to investigate the neural networks underpinning these. Such studies have revealed differences in executive attention networks across cultures (Arora et al., 2020), possibly indicating that this type of attention may be particularly influenced by children's early experiences. Executive attention, however, may not be a unitary contstruct, involving skills relating to planning, decision-making, inhibition, error detection and control of attenion. Indeed, in an update to their original model, Petersen & Posner (2012) suggested there may be two separate neural networks underpinning executive control, though it is possible these become more distinct later in life. It may therefore be helpful to have tasks which can tap different executive skills, such that understanding can be built around how these are potentially differentially related to development. Though a child version of the ANT (the C-ANT) has been developed and can be used with children as young as three (Conejero & Rueda, 2020), this is likely too complex for use with infants and very young children. Tasks which are able to separately measure executve skills in younger ages could thus open doors to build knowledge about the executive attention very early in life, which may be particularly helpful given that this may be a domain particularly impacted by early adversity (Blair & Raver, 2012).

Theoretical work into the structure of attention seems to indicate existence of separate systems relating to orienting, alerting and executive attention, though it remains unclear how social attention relates to these. There is evidence that attention abilities undergo significant development during the first years of life, with a shift from reliance upon exogenous-driven and non-specific orienting systems to more complex executive attention processes (Conejero & Rueda, 2020). Nonetheless, questions remain about how specific executive skills development during this period.

1.3.2 SES and attention research

Given that attention skills develop rapidly during the first years of life, at a time when environmental factors may be particularly influential on a child's development, it seems plausible that early environmental experiences (such as those indexed by SES) may relate to attention abilities. Furthermore, as various work has indicated the importance of attention for other aspects of children's development, it is possible that visual attention is a domain via which SES and cognitive ability are related. In particular, executive attention is likely closely connected to executive functions, therefore consideration of links between SES and visual attention may be particularly insightful for understanding how associations between SES and executive functions occur. This section will outline some of the existing work into associations between SES and visual attention, with some additional discussion of relations between SES and executive functions.

Given the strong focus of developmental literature on the development of social attention, it is perhaps surprising that more work has not investigated relations between social attention and early experiences. Nonetheless, in one paradigm which included a social condition, it was found that high SES (determined as maternal education involving at least one year of college education) infants showed more attention to people and toys during a free play task than low SES peers, with low SES infants showing more inattention (quiet disengagement or not looking) behaviours (Clearfield & Jedd, 2013). In addition, high SES infants showed greater increases in attention when stimuli became more complex. Strikingly, in this study, results indicated that children from low SES families had lower attention abilities across measures of focussed attention and disengagement compared to their higher SES peers already at 6-months-old. Indeed, other work has reported SES-related differences in joint attention that may be apparent already in the first year of life (Reilly et al., 2021). Joint attention is one aspect of social attention for which there has been some investigation of the relation to socioeconomic factors, with some findings indicating different patterns of joint attention relating to SES (Abels & Hutman, 2015; Reilly et al., 2021).

In work which considered more non-social aspects of attention, children aged around 6-years-old from more advantaged families (determined based upon a composite measure of the primary caregiver's education, occupation and income) showed better accuracy attending to a visual cue and were more able to resist the interference of incongruent cues, but did not show differences in orienting, when compared to less advantaged children in an attention network task (Mezzacappa, 2004).This work indicated that SES may be related to alerting and executive attention, but not to orienting abilities. Further support for this comes from the work of Markant et al. (2016) who found a relation between SES (income-to-needs ratio based on education, occupation and income) and working memory skills, but not between SES and orienting efficacy in a sample of nine-month-olds. Such findings suggesting existence of a link between SES and executive skills are not unsupported, with Lipina et al (2005) having found that performance on an A-not-B task differed between 6- to 14-months-old from homes with satisfied and unsatisfied basic needs. Specifically, infants from poorer homes made more errors and performed fewer consecutive correct responses than infants from non-poor homes, indicating differing executive attention abilities. In children aged 10- to 12- year-old, selective attention (the ability to focus on a stimulus in the presence of distracting information) and switch abilities were found to be lower for children from low versus high SES backgrounds (Xia et al., 2024), whilst in 9- to 12-month-old infants, SES was negatively correlated with attention flexibility ability (Conejero & Rueda, 2018). In addition, to behavioural findings, a considerable body of neural research supports a relation between SES and executive attention (see Hackman et al., 2010; Pakulak et al., 2018; Ursache & Noble, 2016 for reviews).

Given the close link between what is considered executive attention and what executive functions entail, it may be helpful to consider relations between SES and executive functions next, such that the potential importance of SES-attention relations may be understood. Curiously, the relation between SES and memory found by Markant et al. (2016) appeared to be moderated by the encoding mechanism used, with an association only appearing when objects were encoded with orienting, and not selective attention, processes (Markant et al., 2016). The authors suggested this may indicate that selective attention could mitigate the impact of SES on memory and may offer a potential mechanism for reducing its impact, thus providing some insight into the link between SES, attention and executive functions.

A substantial body of research also supports a relation between SES and executive functions (EF) (Lawson et al., 2018; Raver et al., 2013). In older children and adults, executive functioning is generally considered to consist of three main domains: working memory, cognitive flexibility and executive attention (Lehto et al., 2003; Miyake et al., 2000). Though this is less clear earlier in development (Friedman & Miyake, 2017; Wiebe et al., 2008), evidence supports that EF skills develop from infancy throughout childhood and later life (Best & Miller, 2010; Diamond, 2002, 2013; Ferguson et al., 2021). It is outside the realms of this thesis to go into great depth here about the development of EF; other work provides more information on this (Best & Miller, 2010; Diamond, 2013; Fiske & Holmboe, 2019); however work investigating SES and EF within the toddler age range (18 months to 3 years) will be briefly summarised.

In a large, longitudinal sample, EF, assessed by performance of three tasks designed to test inhibitory control, working memory and attention shifting, was found to associate with measures of SES at three years (Blair et al., 2011). This finding was replicated in a sample involving participants aged from 3 to 21 years (mean age 12 years) (Ursache et al., 2016). In many of the following studies, the association between SES and EF in toddlers was not the main focus, however other work has found associations between SES and executive functioning, as measured by a dimensional change card sort task, in children aged 3 to 6 years (Henning et al., 2011), between SES and impulse control at 3 years (Bernier et al., 2012; Matte-Gagné & Bernier, 2011) and between SES and EF measured by a battery of tasks from 3 years (Romeo et al., 2022). Somewhat contrastingly, one study found that executive control (as a measure of EF) did not vary between a high-SES and low-SES group at around 3 years (Wiebe et al., 2008). In 2- to 3-year-old children, SES was found to be associated with researcher-rated self-regulation but not with other measures of EF (Elliott et al., 2022).

Other supports of a relation between SES and executive functioning skills comes from slightly older children. At ages six, nine and twelve months, high SES infants (determined as maternal education involving at least one year of college education) showed expected developmental improvements in cognitive flexibility, whereas low SES infants showed a delayed response (Clearfield & Niman, 2012). Noble and colleagues (2007) further found that SES (measured via an income-to-needs ratio based upon education, occupation, and income) explained a significant portion of the variance in visuospatial skills, along with other neurocognitive abilities such as working memory and cognitive control. Though EF is generally considered a key domain which relates to SES, as demonstrated here and in other reviews (Lawson et al., 2018), there is only limited work investigating this relationship in toddlers. In addition, the relationship between SES and executive skills remains unclear and warrants further investigation.

1.4 BRAIN FUNCTIONING

Whilst considering relations between experiences and development at the cognitive level are useful, understanding how brain function links into this could also be of use. The issue of SES disparities is fundamentally a societal issue (Farah, 2017a, 2017b) caused by inequalities within human cultures, therefore some may wonder about the use of a neurocognitive approach to studying it. Indeed, psychology and sociological approaches contribute greatly to our understanding of how children's experiences contribute to their development and outcomes throughout life and should not be replaced by neuroscientific approaches. Nonetheless, considering the association between SES and development from a DCN perspective may provide insights which other methods cannot. Potential advantages include an improved understanding about how SES and development are related, both because neuroimaging may provide a level of measurement and reveal subtle differences which cognitive methods cannot, and because it may reveal mechanistic information about neural processes underpinning the behavioural performances observed. Such insights into how SES-related differences emerge and develop could ultimately inform approaches to ameliorate effects of SES disadvantage. Some advantages of applying a neuroscientific approach to this area of work will be discussed here, using examples of findings which demonstrate these advantages, before summarising the limited literature which investigates the association between SES and brain development in toddlers.

1.4.1 Advantages of neuroscientific methods

1.4.1.1 Earlier detection of SES-development relations

One advantage of neuroscientific methods is that they may reveal impacts of SES which are not (yet) evident at other levels of observation (i.e. in behavioural measures). Indeed, various work has found SES-related differences in neuroimaging measures even where performance levels did not differ. In an fMRI study to compare brain activation during reading-related tasks across children from different backgrounds, Noble et al. (2006) found that 7-year-old children from different backgrounds did not differ in their performance on phonological tasks, but the association between task performance and brain activity did differ according to SES. Phonological task performance predicted brain activation more accurately for children from lower SES families than for children from higher SES backgrounds; specifically, low phonological skill in high SES children was still associated with higher levels of brain activity as it was for children with high phonological skills, whereas low phonological skill in low SES children was related to lower levels of brain activity.

Similar findings have been found using different neuroimaging methods in relation to different aspects of cognition. Kishiyama et al. (2009) found analogous performance levels between high-SES and low-SES 7- to 12-year-olds on a target detection task, however EEG analyses revealed lower ERP amplitudes for the low-SES group. In children aged 12 to 13 years, similar performance levels were found on selective auditory attention task across groups, whilst ERPs showed greater levels of differentiation between relevant and irrelevant tones for the high-SES compared to the low-SES group (D'Angiulli et al., 2008; D'angiulli et al., 2012). Similar findings are reported in younger children, with Stevens et al. (2009) reporting differences in ERP responses during a selective auditory attention paradigm despite comparable behavioural performance between a high-SES and low-SES group (determined based on maternal education) of children aged 3-8 years. The same pattern of findings has also been reported using fNIRs in relation to cognitive shifting task (Moriguchi & Shinohara, 2019). Collectively these findings show how SES may modulate the relationship between brain and behaviour and demonstrate how a neuroimaging approach might reveal SES-related differences which other methods cannot.

1.4.1.2 Mechanistic insights

Exploring mechanistic questions about SES is another potential opportunity facilitated by neurocognitive measures. Whilst behavioural methods can provide some insights about how SES and development are related, neuroimaging methods may uncover information about how behaviours are achieved. In the examples just discussed, it is possible that differences in neural processing despite similar behavioural performance are indicative of the deployment of different mechanisms to achieve a similar result. Other studies in which differing behavioural and neural measures are found could also be informative about the mechanisms by which children are completing tasks. In an fMRI study with 8- to 12-year-olds, both behavioural and neural differences during a stimulus-response mapping task were found between high-SES and low-SES groups, though the manner of these differences was not analogous across the two types of measurement (Sheridan et al., 2012). Instead, high- and low-SES groups showed inverse associations between brain activity and behaviour, which could be due to the use of different neural mechanisms or strategies underlying performance. Studies in domains such as working memory, reasoning and arithmetic processing have also found different patterns between neural and behavioural measures according to SES level (Demir et al., 2015; Finn et al., 2017; Leonard et al., 2019). Whilst neuroscientific research might thus be enlightening about relations between SES and development, it should be noted that they cannot solely inform us about the mechanisms underpinning SES and its relation to brain and behavioural development.

The majority of current work in this area measures associations between measures of SES and development, meaning it is difficult to make conclusions about causal relations. That is, such work cannot identify whether SES factors causally impact brain and subsequently cognitive development, whether neural or cognitive functioning in some way drive SES. Social causation and social selection have both been proposed as two general explanations of this association, though it in fact seems likely that both processes occur (Farah, 2017). Social causation supposes that environmental factors directly influence brain development, which in turn cause SES-related differences, whilst social selection assumes that brain differences are the result of genetic differences, and that these cause SES differences through differences in inherited abilities. Whilst neurophysiological properties of the brain (i.e. heightened plasticity) display how direct SES-brain pathways may occur, findings also support a role for genetics. For example, one study found that the heritability of young children's IQ was modified by SES, such that IQ for the high SES group was predominantly explained by genetic, whereas most of the variance in IQ for the low SES group was explained by environment (Turkheimer et al., 2003). These results support an interaction between genetic and environmental factors in the development of SES-related abilities, which is further supported by epigenetic findings which show a link between environmental experiences and gene expression relating to cognitive skills (Gräff & Mansuy, 2008). Despite practical and moral difficulties around experimentally manipulating conditions of SES, use of random interventions could prove valuable in understanding causal relations between SES, brain, and cognitive development. In one example intervention study, infants of mothers who received a high value monthly cash transfer showed higher beta, alpha and gamma power compared with a control group of infants whose mothers received only a low value transfer at approximately 1-year-old (Troller-Renfree et al., 2022). The authors suggest this offers evidence that an intervention aimed at reducing poverty could cause changes in neural functioning which might be linked to improved cognitive ability, thereby demonstrating the potential insights which interventions studies might provide. Future work would additionally benefit from a longitudinal design, such that time-specific effects can be detected (Fearon, 2019), and must include numerous neurocognitive measures if we are to fully unpick the mechanisms which underpin SES-development relations.

1.4.1.3 Specificity of SES-development relations

As well as affording other advantages, neuroscientific methods may be informative about the specific elements of certain SES environments which relate to development, and about which neurocognitive domains are particularly sensitive to environmental effects.

Much research discussed thus far reveals SES-brain relations largely in language and executive function domains; indeed, in their review of neural correlates of SES in early childhood, Olson et al. (2021) summarise that these are among the neurocognitive areas often found to be particularly related to SES. Specific findings come from work measuring structural and functional brain development, though it is only when studies are reviewed together that specific patterns of relation emerge. These studies are not reviewed here as many are included in the following section, but it is clear that neuroscientific work is adding to knowledge about SES-development relations. As discussed in the previous section here and summarised by others, EF and language are also generally shown by behavioural methods to have the strongest associations with SES (Farah, 2017a). That similar relations are observed at both the behavioural and neural level may be reassuring rather than revealing of novel information about the association between SES and development, whilst reviewing significant and null findings together could also be informative about which neural systems are (or are not) closely related to SES. In particular, future neuroscientific work might investigate associations between SES and a number of neurocognitive domains within one study, such that relative relations can be considered.

Similar consideration to the patterns of findings and relation to the SES measures used could be also helpful in increasing our understanding about which aspects of SES are important for certain areas of development. There is some evidence from both human and animal work that factors such as nutrition, parenting style and stress, among others, are environmental characteristics that play a particularly important role in brain differences (Farah, 2017a). Though such factors may be associated, they are not themselves typically considered measures of SES, but have instead been proposed as proximal factors which may play a mediating role in mechanistic pathways between SES and development (Farah, 2017a). Despite this difference in conceptualisation, investigating how specific measures of SES (such as household income, household CHAOS, etc.) individually relate to neural development in similar ways to how factors such as stress and nutrition have been linked to development, may aid understanding of associations and mechanisms.

1.4.1.4 Electroencephalography (EEG)

Given the possible advantages neuroimaging methods may provide, an increasing number of researchers have begun investigating relations between SES and the brain. This includes work measuring both brain structure and functioning, with work having typically utilised methods such as (f)MRI, electroencephalography (EEG) and fNIRS. This thesis focussed primarily on EEG as a neuroimaging method, as this is a well-established method which holds great potential for more portable data collection, and there is some existing evidence of a link between EEG measures and each of SES and attentional processes, though this is currently understudied. This section will first outline what EEG is, before discussing existing literature which has looked at the link between EEG measures and both SES and attention.

EEG is a well-established neuroimaging method which measures electrical activity from the brain. Electrodes are placed on the scalp, where they detect electrical energy generated by neurons in the form of action potentials, which are how the brain transfers signals. Action potentials are not measured from individual neurons in EEG, however; instead, electrical activity is summed over thousands of neurons to provide an EEG signal. Given that EEG records scalp-level activity, it is difficult to ascertain the source of detected signals, meaning that temporal resolution of EEG is relatively poor. Much EEG work therefore considers activity over general brain areas, such as frontal or posterior regions, and caution should be taken around strong spatial conclusions. Similar to with fMRI, common EEG measures can be broadly broken down into task-related or steady state measures, though the distinction here is a little different. Steady state EEG measures are typically continuous measures (sometimes termed ‘time series analyses’) which may be taken during periods of ‘rest’ but are also commonly collected during exposure to specific stimuli (i.e. social or non-social videos). In this way, steady state might be distinguished from resting-state EEG, which both differ from time-locked task-related measures (i.e. event-related potentials). Event-related potentials (ERPs) are a common measure of neural activity derived from the EEG signal, which consider task-related activity over a very short period. ERPs thus provide information about patterns of brain activity which occur systematically in response to specific stimuli or events. They make use of the high temporal resolution of EEG and require data from numerous repetitions of the same event to reduce the impact of noise and detect a true neural response. ERP measures are heavily used in DCN (S. Morales & Bowers, 2022), with measures of EEG power the next most common. EEG power is calculated by first fragmenting the EEG signal into separate frequency waves, which is generally done using the Fourier transform (S. Morales & Bowers, 2022). Frequencies within a particular range are commonly averaged, leading to measures of specific frequency bands; namely, alpha, beta, theta, and gamma. Power is calculated as the square of wave amplitude and is commonly averaged over multiple electrodes and multiple frequencies within a frequency band. In addition to ERPs and EEG power, there is increasing interest in time frequency analyses, which can be informative about connectivity across brain regions.

1.4.2 SES and brain function research

Whilst much of the existing work in this area has been conducted on older individuals, some work has investigated whether there is a relation between SES and brain functioning in younger children. The brain is especially sensitive to environmental experiences in the first few years of life, in part due to the particularly plastic nature of the brain in this period (Kolb & Gibb, 2011). Though not necessarily critical periods, brain development in the early years is characterised by sensitive periods in which development is particularly reliant upon and influenced by experiences (Tierney & Nelson, 2009). There is thus clear motivation for considering how SES is related to the brain during the first years of life.

Work looking at relations between SES and EEG measures during infancy have revealed somewhat mixed results. In 6- to 9-month-olds, Tomalski et al. (2013) found that frontal gamma power during resting state was significantly lower for infants from low SES families (determined by family income), however EEG resting brain state in newborn infants (aged between 12-96 hours old) was not associated with SES measures (which were parental education, family income and income-to-needs ratio) collected when participants were 15 months (Brito et al., 2016). This work indicates there may be a relation between SES and neural functioning which may only become apparent when children are older than a few days old.

In early work on this topic, Otero and colleagues found differences in EEG power across different frequency bands between groups of children categorised as low- or high-risk according to sociodemographic information at age 18 to 30 months (Otero, 1994) and at 4-years-old (Otero, 1997). More recently, researchers found that SES (measured by a composite of parental education, occupation and income-to-needs ratio) was related to differences in theta power and ERP amplitude in toddlers aged from 16 to 18 months (Conejero et al., 2018). Interestingly, both theta power and the ERP (which was error-related negativity; ERN) in this study were found to be reflective of executive attention demands, thereby the authors suggest these results may indicate how SES could influence early development of the executive attention network. Relations between SES (indicated by a composite of parental education and occupation) and ERN have also been found in 3.5-year-olds (Brooker, 2018). In a sample of 36-month-olds in Bangladesh, higher absolute beta and gamma power in parietal, central and frontal regions whilst children were watching an abstract screensaver and listening to soothing sounds was associated with lower household wealth. Such work supports a link between SES and neural functioning during a child's early years, with indication that this association may be particularly relevant to attention-related neural measures.

Further evidence of this occurs within the auditory attention domain. Differences in ERP response to attended versus unattended stimuli were found for children from higher, not lower, SES families at age four (Hampton Wray et al., 2017), indicating that selective auditory attention appears to be related to SES. In this work, families were specifically recruited from low and high SES backgrounds, where low SES families attended Head Start preschools and mothers of high SES children had a minimum education level involving college education. Age-related analyses further indicated that selective auditory attention was delayed and followed a different developmental pattern for the low SES compared to high SES group. Other neural work supports an impact of SES on auditory selective attention in cases where behavioural performance did not differ, with reduced ERP responses for low SES versus high SES groups when SES is based on various measures including a combination of caregiver education, family income and income-to-needs ratio (Kishiyama et al., 2009), a composite of parental occupation, education and measures of residential area quality (D'angiulli et al., 2012) and maternal education only (Stevens et al., 2009). SES (based on income-to-needs ratio) has also been associated with an ERP measure (the P3b component) relating to inhibition and attention allocation in 4.5- to 5.5-year-olds (St. John et al., 2019). In this study, higher P3b amplitudes were related to higher income-to-needs ratio, but not maternal education, on both go and no-go trials of a go-no-go task. Such evidence supports a particular link between children's early environmental experiences and neural measures related to executive attention.

Within this research, EEG activity within the alpha and theta frequency bands are beginning to emerge as measures that are especially linked to early experiences. In the earlier mentioned study, Otero (1997) found that children categorised as at higher risk of sociocultural disadvantage had greater absolute frontal theta power and lower absolute occipital alpha power compared to children considered not at risk, whilst Jensen et al. (2021) found that maternal perceived stress was positively associated with frontal and central theta power at 36-months-old, though no relations were found between environmental or psychosocial factors and EEG oscillations at 6 months. This relation between stress and altered neural functioning is supported by other work which also indicates different contributions of low and high frequency power which seem to relate to experience of stressful environments. Whilst research has indicated that the relative contribution of low frequency EEG power typically decreases whilst high frequency power increases with age (Marshall et al., 2002), there is some evidence that this develops differently in children who have experienced stress. For example, in a follow-up of the sample of Mexican children used in Otero (1994, 1997), researchers found that differences in theta and alpha power between high and low SES children lessened over age but persisted at 6-years-old (Otero et al., 2003). Specifically, even at this age lower SES children had higher frontal theta absolute and relative power, and lower posterior alpha power (absolute and relative) than children from higher SES backgrounds (Maguire & Schneider, 2019; McLaughlin et al., 2010), indicating that typical maturation patterns of decreasing lower power and increasing higher power may be different in these children. Such findings are also supported by other work which found smaller changes of relative theta power over age in children who had experienced stress or disadvantage compared to children who had not (Harmony et al., 1988) and by associations between experience of institutionalisation and EEG power (Marshall et al., 2004, 2008). Whilst many of these studies used parent-report or environmental measures relating to stress, there is also evidence of relations between stress and neural functioning when measures of physiological stress have been used. Troller-Renfree et al. (2020) used hair cortisol as an index of maternal stress, which they found was related to relative theta and alpha power in 6- to 12-month-old infants. In line with previous work, higher physiological stress was found to relate to higher theta and lower alpha power, thereby indicating different distributions of low and high frequency power across development in relation to early experiences.

1.4.2.1 Why alpha and theta frequencies specifically?

Activity in the alpha and theta frequency bands may be particularly relevant here as, in addition to findings just discussed, they are commonly implicated in attention processing. There is various evidence that activity in the theta frequency may be particularly related to executive attention, whilst alpha activity has been implicated in relation to inhibitory control.

Whilst theta activity has been linked to a variety of cognitive processes, evidence has emerged that frontal theta may be particularly related to inhibitory control (Eisma et al., 2021; van Noordt et al., 2022), with some researchers having proposed it as a candidate mechanism for this (Cavanagh & Frank, 2014). Other evidence indicates a relation between frontal theta and each of error detection (Conejero et al., 2018), and conflict management (Jiang et al., 2018; Kaiser et al., 2022), which both related to executive skills and thus further support this link between frontal theta and executive attention. Additionally, another line of work has indicated a role for hippocampal theta oscillations in working memory processes, with some suggestion that these two theta systems may interact with one another (see Senoussi et al., 2022 for discussion). Whilst theta seems to relate to a variety of executive skills, the role of alpha activity appears more closely link to inhibition specifically (Klimesch et al., 2007). Various work has indicated a reduction in alpha activity contexts where participants observe or executive goal-directed behaviour, in which irrelevant information must be inhibited (Nyström et al., 2011; Southgate et al., 2009). Parietal alpha has also been found to relate to focussed periods of attention to the centre of the visual field (Klimesch et al., 1999, p. 199; Orekhova et al., 2001; Westphal et al., 1993), which might further support that alpha is implicated in inhibitory controla dn maintenance of focussed attention (Orekhova et al., 2001).

In addition to these specific findings, there are also more general indications of links between alpha and theta, and attention. For instance, various work indicates the existence of altered alpha and theta oscillations related to attention deficit hyperactivity disorder (ADHD) (Cross-Villasana et al., 2015; Guo et al., 2020; Lenartowicz et al., 2014; Mann et al., 1992). Since attention is also altered in ADHD, such findings could support a possible relation between activity in these frequency bands and attention. In addition, theta and alpha power are the frequency bands most implicated in social processing (van der Velde et al., 2021), with various work indicating that theta and alpha power are different in conditions which involve social versus non-social stimuli (Jones et al., 2015; Orekhova et al., 2001; Stroganova et al., 1999). Given that processing of social stimuli constitutes social attention and may place greater attention demands that simpler stimuli, this association may be due to the involvement of alpha and theta power in attention processes. Work that found relations between alpha and theta activity and attention where these were not apparent in other frequencies (i.e. Xie et al., 2018), support that these two frequency bands may be particularly important for attention processing and subsequently may be especially relevant when considering relations with early environmental experiences.

1.4.2.2 Conclusion

In conclusion, these findings contribute to growing indications that children's early experiences may be related to neural functioning in the theta and alpha bands particularly, with work suggesting that theta power in frontal regions and alpha power in occipital regions may be especially relevant. Given evidence that alterations in brain development such as these can be long-lasting impacts on cognitive development, there is strong motivation to better understand how these differences can occur such that effective support can be developed where necessary. Though there is an increasing amount of work focussing on better understanding SES-brain associations in early development, only a small amount of this work to date has involved toddlers. In fact, in a review of SES and brain research during early childhood, Olson et al. (2021) found just nineteen research papers which fulfilled their criteria including participants aged 5 years and under. These nineteen papers included work using any brain imaging method, meaning only a small number of these were actually focussed on relations between socioeconomic experiences and theta or alpha power. There is thus a clear need for work which focusses on neural development in the alpha and theta frequency ranges during the first few years of life, such that we can build greater understanding about how this is altered in relation to SES.

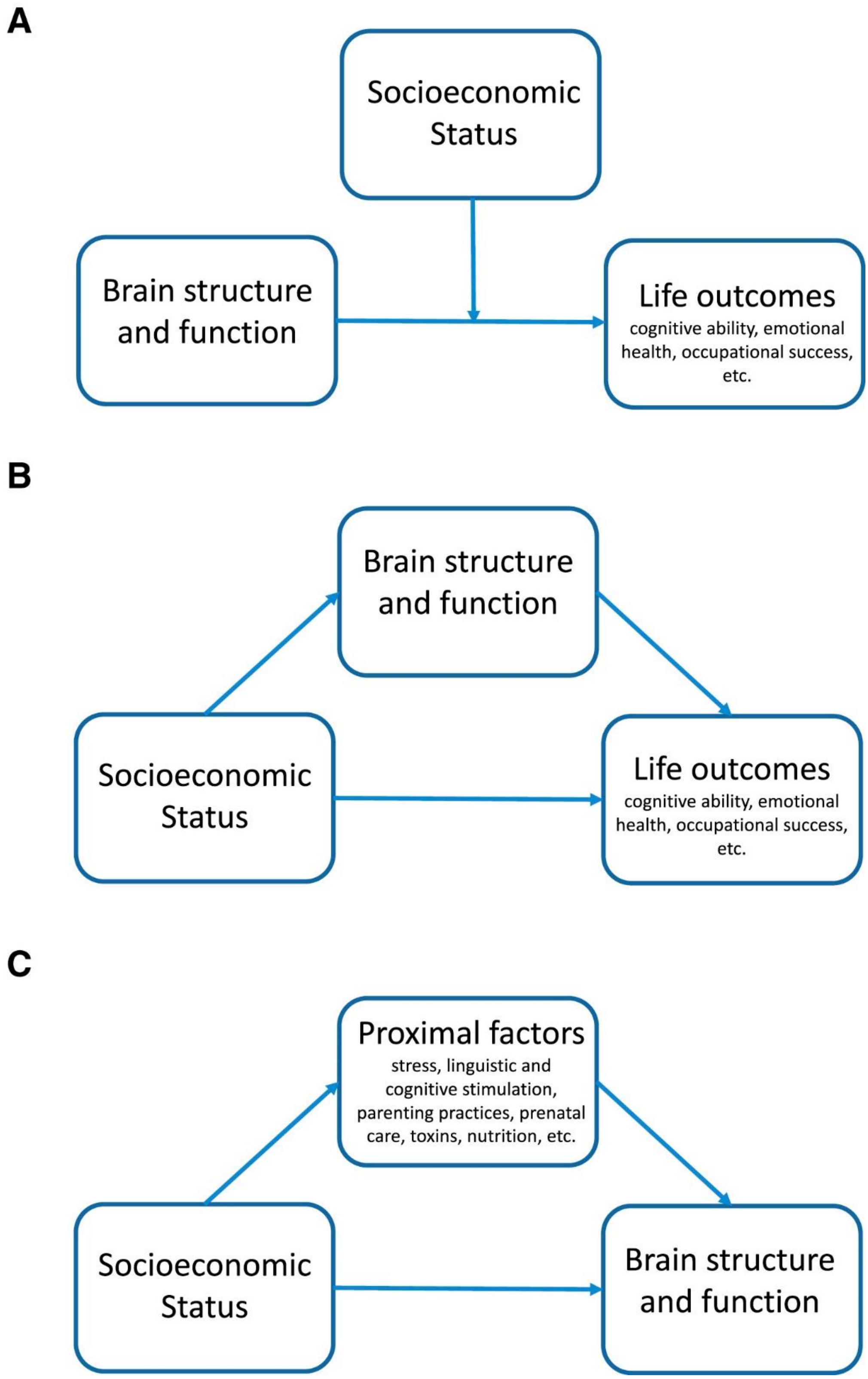
1.5 THEORIES ABOUT CONNECTIONS BETWEEN SES, ATTENTION AND BRAIN FUNCTIONING

The previous sections of this chapter have outlined evidence of relations between SES, attention and neural functioning, but have not discussed how/ why these domains may be related. Such theoretical considerations are important as they provide a framework of understanding which brings together existing findings and guide future research in this area. The current section outlines some key theoretical ideas about connections between these domains.

In their review about the neuroscience of SES, Farah (2017) outline three possible models explaining relations between SES, brain function and behavioural outcomes (Figure 1.1). The first model demonstrates a moderating effect of SES on the relation between neural functioning and cognitive outcomes such that children from low versus high SES backgrounds engage different patterns of neural activity to perform the same cognitive skill. As Farah (2017) discusses, there is evidence that performance in tasks demanding executive attention skills may be underpinned by different neural functions in children from different SES backgrounds, despite no observable differences in behaviours (D'Angiulli et al., 2008; D'angiulli et al., 2012; Stevens et al., 2009). The second model outlines the brain as a mediator in the relation between SES and behavioural outcomes. In this model, SES-related neural differences could, but do not necessarily, predict SES-related behavioural differences. Some existing work has used mediation analysis to investigate this kind of model, with Noble et al. (2015) having found that SES predicted cognitive ability as well as structural brain measures, and neural structure accounted for a significant proportion of the relation between SES and executive function. Such findings may support that SES relates to cognitive abilities associated with executive attention both directly and indirectly via effects on the brain. Finally, the third model suggests that proximal factors associated with SES (i.e. stress or parenting) may mediate associations between SES and the brain. This model assumes a causal relation between SES and the brain, which may occur either directly or via relations to proximal measures. Research has again benefitted from mediation analysis in observational studies to investigate this question, with various work indicating that measures of stress appear to mediate relations between measures of SES and neural activity (Kim et al., 2013, 2016; Luby et al., 2013). In the context of executive skills, maternal parenting behaviours and parenting stress have been theorised as pathways via which SES may have an impact (Vrantsidis et al., 2020), though in a large sample (N = 2214) of young children, only maternal cognitive stimulation and not harsh parenting and parental stress emerged as an indirect pathway, with evidence also of a direct link between SES and sustained attention (Meng et al., 2024). These differences may be due to demographic or measurement differences between the studies, and could indicate that different models and relations may be at-play across development.

Indeed, whilst Farah's (2017) third model assumes a causal relation between SES and development (which can be broadly termed ‘social causation’), it is likely that another process (sometimes termed ‘social selection’) may also be in operation. Social selection purports that genetic differences cause neural differences, which in turn cause different educational and occupational performances, thereby leading to differences in SES. In the context of development, childhood SES (i.e. that children are raised in the parents’ SES contexts) is thought to relate to neural and behavioural phenotypes because children inherit genetic predisposition to SES levels akin to that of their parents. In this way, social selection considers that neural differences are under genetic control and may lead to, rather than being caused by, SES differences. This issue of causality is tricky to unpick. Whilst many people might intuitively assume that SES impacts attention and brain development, there is little work that is able to actually investigate this causal relation. Despite ethical and practical difficulties with implementing manipulations to socioeconomic factors, however, there is now one research study that has implemented a randomised control trial of poverty reduction on children’s early development (Home | Baby's First Years (babysfirstyears.com)). Findings from this study have indicated that providing families with high-value monthly cash donations led to differences in infants’ high frequency power at 12-months-old which were not apparent in a infants whose parents received a low-cash monthly donation (Troller-Renfree et al., 2020). This finding suggests that changes in family income may lead to changes in neural development, though considerably more research is needed to fully elucidate this relation.

Figure 1.1: Figure taken from Farah (2017) showing possible relations among the causes and consequences of SES and its neural correlates



1.6 GAPS/ LIMITATIONS IN EXISTING KNOWLEDGE AND LITERATURE

A DCN approach to the relation between SES and development has many advantages over more traditional approaches, primarily holding potential to provide insights about *how* SES and development are related, including about mechanisms which may underpin this association. Existing research indicates activity within the theta and alpha bands may be especially linked to SES and attentional processing, yet there is potential for much future research to further elucidate specifics of these associations. The limited research that exists, however, tends to include children aged 3 years upwards and may not hold when looking more specifically at the toddler age range.

1.6.1 Toddlerhood as a crucial but challenging time in development

As indicated by the work reviewed so far, a huge amount of brain development occurs in the early years of life which may be crucial for functioning throughout life, thereby signifying it as an important period for developmental cognitive neuroscience research to focus on. This thesis primarily includes work with toddlers, which is defined here as the ages between 18-months-old to 3-years-old. Whilst there is already a great deal of developmental cognitive neuroscience work investigating infants (<18-months-old) and older children (typically from 3 years upwards), research involving toddlers is generally lacking. This may be, at least in part, due to practical difficulties with including toddlers in neurocognitive research, difficulties which arise due to the development which occurs at this age. This section will outline the importance of including toddlers in DCN research, before explaining some difficulties of doing so and briefly suggesting some methods which may be useful in broadening research to this age range.

At around 2-years-old children undergo massive increases in their language, motor and cognitive abilities (Calkins, 2007; Colson & Dworkin, 1997) and are on the cusp of considerable executive function (EF) development (Anderson & Reidy, 2012). Rapid social and emotional development also occurs during toddlerhood, with the emergence of self-awareness and development of self-regulatory skills (Kopp, 1982). Findings support that the development which occurs during toddlerhood is crucial for later development, with research indicating that the rate of learning, as well as measures of skill, over this period may be associated with later ability. For example, the velocity and acceleration in vocabulary growth between 14 to 46 months has been found to predict vocabulary at 54 months (Rowe et al., 2012), whilst working memory measured at around two and a half years may be predictive of classroom engagement, receptive vocabulary and number knowledge at age six (Fitzpatrick & Pagani, 2012). Development in toddlerhood may lay foundations for much later in childhood, with evidence that language skills at around 33 months were associated with language ability at around 10-years-old (Reikerås & Dahle, 2022) and ability to restrain oneself between 14- to 36-months-old may be important for EF skills at age 17 years (Friedman et al., 2011), among others. In sum, SES-related differences in neurocognitive functioning are already apparent by preschool age (Janus & Duku, 2007) and the brain may be particularly susceptible to environmental input early in development (Merz et al., 2019), indicating the importance of investigating development during toddlerhood.

Though this period of significant change marks toddlerhood as an incredibly interesting age to study, it also introduce many practical challenges for data collection. Though numerous neurocognitive longitudinal studies now follow children through toddlerhood, there remains a deficit of research into this age range, at least in part due to practical difficulties associated with testing toddlers (Hendry et al., 2016). Motor developments typically mean toddlers are recently mobile and motivated to physically explore their environment, though their attention skills do not enable them to focus on one thing for long. Despite improved language abilities compared to infants, these skills are still very much developing throughout toddlerhood and there may be large variability in children's understanding of an experimental task, meaning language responses cannot be relied upon entirely. It can thus be difficult to gather enough meaningful data from toddlers for analysis and there is clear motivation for work which focusses on measuring development during toddlerhood, particularly in children at risk for poorer outcomes.

When researching a period of rapid development, collecting data in a timely and effective way is critical. Though some studies conduct lab visits as often as monthly, development occurs more rapidly than this, meaning current methods may miss some details of development (Fearon, 2019). Furthermore, measures collected during lab visits may be influenced by variable factors such as mood, sleep, and time of day. Current behavioural methods tend to use standardised assessments which can be long and tiresome and may reflect children's mood and engagement levels as much as their cognitive ability. There is thus a need for measures which capture children's true cognitive capacity effectively. Portable methods (smartphone, wearable devices, etc.) enable data to be collected regularly from community or home environments and may enable some of these difficulties to be overcome (Fearon, 2019). Individualisation of methods can additionally limit the effects of mood and engagement on measures, by ensuring the data collection experience is accessible as possible for each participant. Individualisation may also enable assessments to be shorter and more efficient in detecting children's cognitive ability and could be a useful tool in future developmental research. Nonetheless, methods must be designed such that data collection allows participants freedom to move and explore their environment and does not rely on language comprehension or expression. Though neuroimaging methods require a certain level of stillness to ensure data quality, technological advancements have facilitated the development of wireless systems, which may present an opportunity for collecting neurocognitive data which may be particularly useful for toddlers.

In conclusion, toddlerhood poses a particularly interesting period during which children undergo considerable neurocognitive development across a range of domains. Whilst indicating the importance of this period, these developments also contribute to significant challenges with including toddlers in DCN research. Some such challenges and how recent methodological advances may provide opportunities to overcome some of these have been briefly discussed here.

1.6.2 Lack of diversity/ representation

In addition to a lack of studies focussing on toddler development, another shortcoming of existing developmental research is that it is under-representative of human populations in various domains, including race, ethnicity and socio-economic background of participants (Henrich et al., 2010) and researchers (Garcini et al., 2022). Given that research thus far indicates that environment-SES relationships are significant, it's remarkable how little variety of experiences the field currently encompasses. This under-representation is driven by an existing literature which is driven largely by work involving samples of Western, white, middle-class children (Rowley & Camacho, 2015) and is likely contributed to by multiple factors, including practical, personal and systematic barriers to participation (Garcini et al., 2022). Despite challenges, it is critical that DCN research increases diversity in its samples for multiple reasons, including the need to increase generalisability of findings and to create effective support for the populations who may benefit the most from early intervention.

1.6.2.1 Why increasing diversity is critical

Firstly, research that includes only a specific set of individuals is less generalisable than that which more accurately reflects the population. Historically, psychological and SCN research has typically included certain groups, leading to them being overrepresented in the literature. In order to overcome this and improve the representation of our research, innovation of different approaches and dedicated efforts are needed. Secondly, as alluded to above, including individuals with a wider diversity of experience may facilitate insights about the impact of different experiences on development which may not be possible without such rich variety. In turn, this could help us to understand the processes and mechanisms through which these effects occur and may enable more effective support to be developed, ultimately ameliorating early negative effects. Finally, this work is also motivated by a cultural purpose. Individuals who are typically less represented in DCN research are also those who have been more commonly prejudiced against. Such persecution has multiple effects, including limited access to support, being less likely to engage with support/ know how to and, in some cases, experiencing worse health care (Garcini et al., 2022). In order to improve and overcome such challenges, work must be done to empower often-marginalised communities and to improve understanding and respect between individuals with different experiences. This is crucial to the purpose of work which seeks to improve developmental processes and outcomes for all individuals.

Simultaneous overrepresentation of one demographic and underrepresentation of others may lead to a skewed interpretation of what constitutes ‘typical’ development and consequently of what is ‘atypical’ (Jordan & Prendella, 2019). By developing support and interventions for low SES families based upon a typical range determined by children from high SES families, it is likely these are not maximally effective and are limited in their impact. In addition, the misrepresentation of research influences how findings are framed, such that potentially positive, adaptive abilities in some children may be considered as negative maladaptation. This is argued by some researchers, who suggest that low income environments cue an adaptation of psychological processes which lead to performance on tasks which are rational in the context of resource scarcity, environmental instability and other SES-related factors, but which may hinder success outside of this environment (Sheehy-Skeffington, 2020).

1.6.2.2 Challenges of improving diversity

Despite work outlining the need for change and ways to do achieve this, underrepresentation of the diversity of socio-demographic, ethnic and racial participants from which study samples are taken remains a substantial issue in developmental research (Garcini et al., 2022; Jordan & Prendella, 2019; Rowley & Camacho, 2015). Review work has shown that an overwhelming majority of psychological and developmental research includes “WEIRD” participants, despite this demographic making up only around 12% of the World’s population (Arnett, 2008; Henrich et al., 2010; Nielsen et al., 2017). “WEIRD” (Western, Educated, Industrialised, Rich and Democratic societies) (Henrich et al., 2010) is the term coined to describe the demographic typically included in DCN research, though it should be noted here that this is just one way of highlighting the clear bias in existing DCN work. Green et al. (2022) choose to outline what they term exclusion bias (exclusion of marginalised groups due to study inclusion criteria), sampling bias (occurring due to recruitment strategies not focussing on diverse groups) and loss of diversity due to attrition of marginalised groups in studies. However it is conceptualised, it is apparent that DCN is severely underrepresented in terms of diversity across a range of domains and active efforts to overcome challenges and improve inclusion of the field may be needed to improve this (Garcini et al., 2022).

A multitude of factors may contribute to this skew. From a practical perspective, many child development studies take place in a lab setting. This mean families must have the time, resources, and motivation to travel to labs and, though many labs do reimburse travel costs, travel logistics could be a significant barrier to lower SES families taking part. Some parents may have no experience of research and not understand the function of research or the potential value of their children taking part in research. For other families, historical oppression and marginalisation by societal organisations may contribute to mistrust of research institutes (Garcini et al., 2022). Related to this are concerns around how participant information will be used, which could perhaps be exacerbated by uncertainty about why diversity-related information (such as ethnicity, SES information, etc) is being gathered.

Whatever the reasons for such unbalanced representation in DCN research, it is imperative that this is improved. Some ways in which this may begin to be achieved will be presented next. Notably, strategies for improvement can be employed at various stages of research including during design, data collection or post-data collection (Garcini et al., 2022 for discussion). Firstly, studies could be designed based on theories which do not discriminate against certain groups of people; for example, by using an adaptive framework which focusses on strengths rather than just deficits. The study design process could also involve participants and community members to ensure that research is effective in supporting the people it aims to reach and to help build a trusting two-way relationship between researchers and the community. This latter benefit might also be helped by a post data collection commitment to continued engagement with the community. A current perspective of some participants might be that their data is taken and used, without any meaningful return for them; this could contribute to mistrust and unwillingness to participate. By continuing to share project progress and communicating other scientific information to communities, long-term relationships which benefit both researchers and communities could be established (Green et al., 2022). Other longer-term strategies for improving diversity in research relate to the continued promotion of diversity in research, in scientific communities and events, as well as within research itself (see Swartz et al., 2019 for recommendations about how to promote diversity in scientific organisations).

Within the realm of data collection, improving recruitment strategies is a key focus. Practical adjustments which researchers could make include offering transport to and from testing sessions, flexible scheduling around families' commitments, free care for siblings and material incentives including monetary compensation, vouchers, and offers of food and drink (Garcini et al., 2022). These modifications may seem minimal to some but may contribute to significantly reduced burdens for participants. In addition to reducing burdens associated with attending a lab setting, development of effective tools for collecting data in community settings may be another way to reduce some of the requirements for families participating in research, enabling more diverse participation. Technological advances in neurocognitive methods (such as eye-tracking, EEG and fNIRs) mean that some systems are transportable and can be taken into community or home settings (Garcini et al., 2022), meaning that research could take place in trusted locations which are familiar to participants. Other tools include applications which can be remotely downloaded and used via participants, without the need to ever meet or interact with researchers, which may be beneficial for some participants. Other approaches researchers could use include improved explanation of the data collection process, such as providing a video explaining the methods used, and using tasks or setups which are sensitive to differences across cultures or groups (Garcini et al., 2022).

Given the dire need for increasing diversity, this thesis is set in a context of work which aims to improve diversity and representation in developmental cognitive neuroscience (DCN) research. Whilst it might be argued that *all* DCN research should work to improve diversity in their research, unfortunately the reality is that this is not a priority of most studies (Garcini et al., 2022), which is why work such as this thesis are needed to be dedicated to this endeavour.

1.6.3 Issue with definitions and measurements of early experiences

1.6.3.1 Challenges of defining and measuring SES

Despite such a wealth of findings relating SES to development, there remains no one clear definition of SES or a single framework which incorporates all the many dimensions of differences which different individuals experience. Indeed, this multifaceted nature of experience is partly what makes it difficult to define and measure. Nonetheless, it is crucial that we develop a better understanding of how early experiences shape children's development to help all children achieve their optimal outcomes in life. This may be particularly important where experiences are negatively impacting development and may help us to better understand developmental differences across cultures and contexts.

The SES-development literature uses a variety of measures of SES, ranging from individual measures of education to composite measures which incorporate numerous dimensions, and including those that measure the area in which a child lives (i.e. the Index of Multiple Deprivation) as well as measures of the home environment (such as the Home observation measure (Caldwell & Bradley, 1979)). In the discussion above, any such measures have been considered as SES, in order to provide an overarching picture of the current field, however it could well be argued that this view is simplistic and incorrect. As purported by some researchers, this variety of measures consider different phenomena which, though often related, may reflect distinct sources (Braveman et al., 2005). By not treating these as individual concepts, we may miss out on informative details about the relation between experiences and development. In turn, deriving a composite score to use as a single measure of SES may result in a measure which inaccurately groups distinct concepts together and is less informative than its composites. Treating measures as individual measures, however, also has limitations – not least due to the practical difficulties of investigating every element of SES, but also determining which of the many factors may be important for each neurocognitive area and how they interact with one another.

Deciding what measure of SES to use is at root a theoretical question, which depends on how SES is defined, however in practice SES is often defined by how it is measured. In fact, a recent literature review found that nearly 80% of psychological research papers about SES did not define SES theoretically (Antonoplis, 2023). This circularity leads to confusion around how to interpret findings; without definitions of SES, it is difficult to interpret SES measures, yet without measures, there is currently very little definition of SES. To overcome these difficulties, Antonoplis (2023) proposes a definition of SES which states that "SES represents individuals' possession of normatively valued social and economic resources" (p17). In addition to providing a definition of SES which is separate to how it is measured, this definition also more clearly sets out what SES is and means, whilst also fitting in practically with the current field of research (see Antonoplis, 2023 for discussion). Following this, the author further suggests a reconceptualised approach to studying SES as a combination of structural features which are each individual indicators of SES. Here, some common difficulties/ decisions which researchers face when determining a measure of SES will be discussed before it is explained how reconceptualising SES in this way may help answer these difficulties.

1.6.3.2 Subjective versus objective SES

Though perhaps the most commonly considered measures of SES are those that are separate from the participant and measure from the outside (i.e. are objective measures), subjective measures of SES can also be gathered. For these, participants will typically be asked to compare themselves in relation to others within a particular society, meaning they are measures of an individual's integration of factors which they consider important for their social standing within a given community. As such, subjective measures may reflect factors which objective researchers may not know to ask about and which objective measures tend not to capture but, by their nature, make comparisons across individuals difficult.

1.6.3.3 SES is dynamic

Although for the large part SES is discussed as though it is a fixed measure, it is in fact dynamic both over time and levels of organisation. That is, SES does not operate only at the individual level, but factors can also impact the household or neighbourhood in which an individual lives (Hackman & Farah, 2009). With regards to time, factors of SES such as income or education change over the course of a lifetime, often slowly but sometimes at a rapid pace. Measuring SES at a single timepoint might therefore be considered only a single snapshot of that person's experiences which do not fully reflect their true experience. This consideration might be particularly pertinent when considering neurocognitive development, as associations between SES and development might have a temporal element such that certain experiences may impact development only if they occur at a particular point in development.

1.6.3.4 Parental versus child measures

In DCN research, SES measures typically relate to parents rather than the child specifically. This is partly due to practical reasons related to how SES is typically measured, though one can also argue that much of a young child's SES is reliant upon their parents' or caregiver's SES (Hackman & Farah, 2009). This may be somewhat related to the previous point about dynamic SES in that it may be difficult to determine at what point measures should change from parent-focussed to child-focussed and therefore what the most appropriate measures are in a given study.

1.6.3.5 Measuring SES: dichotomous or gradient

Another consideration when choosing a measure of SES is whether measures should be dichotomous or continuous. Existing SES research has shown that SES-related disparities (including those at the neural level) seem to occur on a gradient over the entire SES range, rather than only occurring at a specific threshold (Duncan & Magnuson, 2012; Raizada et al., 2008; Spann et al., 2020). Such results suggest that a continuous measure of SES may be most appropriate when further investigating this relation, though research has, on occasion, found that SES disparities were only apparent when categorical comparisons were conducted (Moriguchi & Shinohara, 2019). To establish the nature of association between each SES component and development, it may be useful to first use *both* a categorical and continuous approach, though once determined, the appropriate measure should be used. It should also be noted that, whereas continuous measures may be more suited to comparisons across different study samples, categorical SES measures which are split into low and high groups based upon the distribution of the study sample are somewhat context-specific and should not be generalised across studies.

1.6.3.6 Defining SES

As is clear from the discussion above, determining a definition and measure of SES is not a simple choice, with no clear guidelines for best practice. A key question about SES is around whether measures should be combined into a single, composite measure of SES or whether individual measures should be kept separate, whilst Antonoplis (2023) makes a convincing argument to reconceptualise SES as a set of features which each contribute to an individual's experiences; it is these features, rather than SES itself, that may be linked to development. One might consider factors such as age, gender and race as features of an individual's environment which impact their status in society ('structural resources'; (Halasz & Kaufman, 2020)). Under Antonoplis' (2023) conceptualisation, SES features should be treated similarly to how age, ethnicity or gender are treated, as these are also properties (though distinct entities) of an individual's environment (as are SES features), and choice of statistical analyses should reflect this. Structural features of SES could include any measure which is thought to relate to an individual's economic standing in society; this might include traditional variables such as parental education, occupation, and household income, as well as other measures of resources, chaos, and other features of the home. Under this framework, certain individual SES features may be more or less likely to lead to other conditions occurring (Antonoplis, 2023).

This reconceptualisation has numerous advantages which make it a convincing proposal. Firstly, reorganising SES in this way moves the focus away from developing a valid measure of SES towards investigating which structural features may be most relevant to the neurocognitive construct being studied (Antonoplis, 2023). In this way, findings will be referred to in relation to those specific features only, rather than being inferred to mean something more broadly about SES as a whole. This may offer a second advantage in that sense may be made of existing findings which use varying measures SES. Studies which measure the same SES features may be considered together and conclusions could be made based upon only those features rather than SES more broadly. In this sense, this conceptualisation of SES does not discount the wealth of already existing research but provides a framework within which it can be interpreted. This focus on more specific features of SES may also be useful for the development of interventions, as these could be more precisely designed and ultimately may result in more effective support. Furthermore, the idea that numerous features or conditions together contribute to SES is in line with other researchers who have defined SES as a ‘complex bundle of factors that are generally but imperfectly correlated’ (p56, Farah, 2017), those who have argued that individual measures of SES factors capture aspects of social indicators of outcome that all-purpose measures cannot (Debrouwere & Rosseel, 2023) or similar (Braveman et al., 2005). Where SES is discussed in this thesis, it is considered under this framework, whereby different measures of an individual’s social standing such as income and education are considered SES features which each contribute to an individual’s experiences.

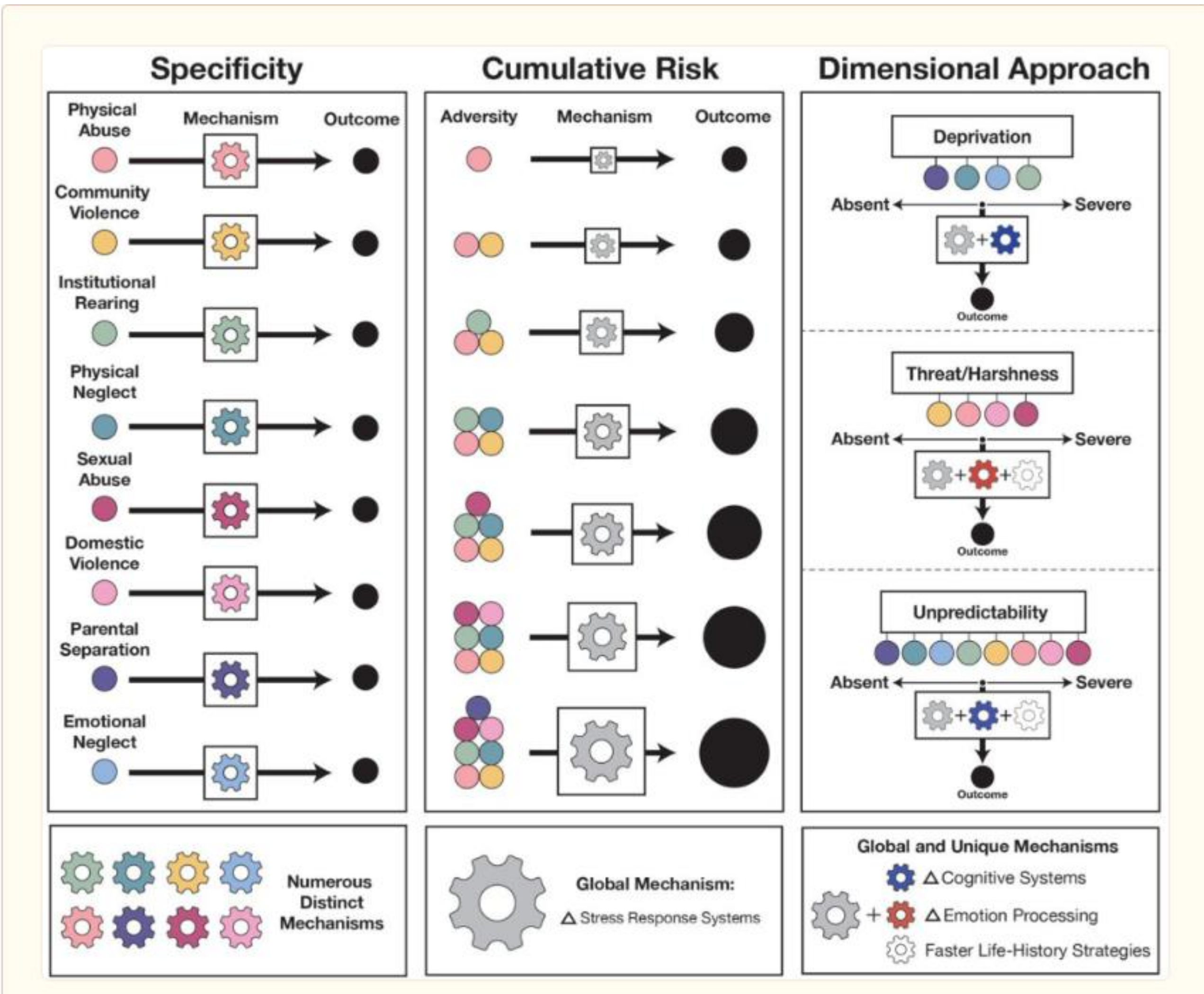
1.6.3.7 Theoretical framework of SES

In addition to challenges around what constitutes SES and how it should be measured, consideration should be given to theoretical frameworks accounting for the relation between children’s experiences and their development. In much of the work about this, an assumption is made that lower SES (i.e. more adverse, stressful conditions) is associated with poorer development; a perspective which largely defines the deficit model of development. This model contradicts an alternative adaptation-based view, which considers that children develop in response to their experiences. As such, children who experience more adversity may appear deficit in certain neurocognitive skills when considered within a low-adversity context, but may be well-adapted to high-adversity contexts (Ellis et al., 2017). This section will further explain the adaptive framework of development, discussing some research which supports it and briefly outlining other, related theories where relevant.

As mentioned, much existing work presents SES-related differences in neurocognitive development as deficits caused by disadvantage, an interpretation which has powerful impacts on how society views children from low-resource backgrounds. Whilst this is supported by various research findings, it can be argued that the deficit model neglects to consider how SES-related differences may be adaptations in neurocognitive abilities which individuals make in response to the conditions of their particular environment (Ellis et al., 2017). It emphasises weaknesses and overlooks potential strengths which arise from experience of low SES conditions – indeed, these strengths have been termed 'hidden talents' by some researchers because they are invisible to much of society (Ellis et al., 2020). Policy-makers, educators and the general public – including individuals themselves – are all impacted by this unbalanced presentation of strengths and weaknesses (Ellis et al., 2020) of children from low SES backgrounds, which can continue to contribute to less-positive experiences for them. Crucially, a deficit-based view may impact the support which children receive, perhaps by promoting remedial rather than growth mindset around their development (Ellis et al., 2020).

A deficit model is central to some mechanistic theories about how SES impacts development. Cumulative risk theory purports that the level of SES disparity depends on the number of adverse experiences that an individual experiences; all experiences add together to create a level of risk or deficit which is equal to the sum of its parts. The underlying assumption is that the higher number of adversities an individual experiences (i.e. the more low SES features), the greater the developmental deficit (Ellis et al., 2017). Dimensional models of adversity are also somewhat based in a deficit model. In these, dimensions of experience are identified and are considered to vary on a spectrum of severity and chronicity; their effect on development is thought to be as a function of their position on this spectrum. Though focussed more on developmental mechanisms involved in this relationship than cumulative risk approaches, dimensional models at root consider that a higher level of severity/ chronicity for a dimension of experience is associated with worse development. Figure 1.1 below shows a visualisation of different theories about how environmental experiences might influence development.

Figure 1.2: Taken from McLaughlin et al. (2021). This image shows three distinct approaches that have been proposed about the mechanisms by which early life experiences might impact development.



1.6.3.7.1 A note on SES and adversity

A note should be made here about the relation between adversity and SES. In addition to confusion around how measures of adversity and SES may be related, there is ongoing debate around a definition for adversity. Whilst some researchers have defined adversity by its relation to stress, issues with this approach have led others to define it as deviations from the expectable environment that may require substantial adaptation (McLaughlin, 2016; McLaughlin et al., 2019; Nelson & Gabard-Durnam, 2020). The expectable environment refers to anything that the human brain expects to experience to develop typically, which might include specific or broader inputs. For example, it is relatively well-established that the eye requires exposure to light so that the visual system can fully develop (Fox et al., 2010), whilst work suggests that experience of a responsive and sensitive caregiver may be important for social development (Nelson et al., 2019). Under this framework, adversity is defined by either the absence (i.e. lack of caregiving, nutrition, etc.) or the presence (i.e. physical abuse) of an unexpected experience which is deemed to have a negative impact on development (Nelson & Gabard-Durnam, 2020). In this way, experiences are somewhat defined by their severity, with only those that are 'severe enough' (p279) or chronic qualifying as adversity (McLaughlin et al., 2019).

A definition of SES, by contrast, typically centres around factors relating to an individual's standing in society, rather than by their effects on development. Whilst measures of SES (i.e. income) may relate to aspects of adversity (such absence of food, clothing, caregiving, exposure to violence), they do not necessarily capture the same thing. One conceptualisation may be to consider that both SES and adversity can share features, but interpretation of these features as measures of adversity might depend upon context or feature characteristics. For example, features such as scarcity of food or resources *could* both be considered features of SES as well as indicators of adversity, though whether they are considered relevant to adversity might depend on their severity and impact on development. It could then be that SES includes less-extreme features that aren't considered adverse as they do not significantly and negatively impact development but could plausibly still have a lesser effect. This framing would consider the impact of environmental experiences on a continuum, which might be supported by work which finds an SES gradient (Noble et al., 2007), but poses questions about how the distinction is made between adverse or not adverse. Alternatively, it may be that SES is a conceptualisation of socioeconomic adversity, whereby SES features such as income, parental education and occupation are thought to have a relation to development which is distinct from other aspects of adversity such as threat or deprivation (Farah, 2018). This may be somewhat supported by a dimensional theory of how environmental experiences relate to development, as this suggests different adversity types might impact development in different ways (McLaughlin et al., 2019), though notably a dimensional approach does not necessitate that all individual adversity types are functionally distinct (McLaughlin et al., 2021). Support for this view comes from findings which show evidence that experience of neglect might affect the brain differently to how threatening events do (Sheridan & McLaughlin, 2014), whilst it seems plausible that the type of each adverse event (i.e. physical or visual) might also differentially impact development (Nelson &

Gabard-Durnam, 2020). Given these differences and under this framework, it has been suggested that findings about how adverse childhood experiences relate to development may not necessarily generalise to poverty or SES, therefore caution should be taken when comparing findings about these (Farah, 2018).

In addition to other ambiguities about how to define SES, consideration about how SES may fit in with the concept of adversity should be considered. As explained here, it might be that SES and adversity share features but are distinguished from one another by their effect on development; alternatively, SES may be considered a type of adversity or there may be another explanation not covered here. Though important to contemplate in a wider understanding about the relation between environmental experiences and development, answering this question is outside the remit of thesis and require greater focus to be better understood.

1.6.3.7.2 Adaptive framework

I have thus far outlined some limitations of a deficit-based model of how experiences impact SES; now support for an adaptive theory will be considered. Important to any conceptualisation of SES-effects is understanding how and why these occur. In addition to societal motivations for considering individuals' neurocognitive strengths as well as weaknesses, this model also gains support from evolutionary-development theories. Within such theories, adaptation is key to survival and reproduction and need not be only beneficial, but benefits must outweigh costs (Ellis et al., 2017). Adaptations are driven by selection pressures of an environment, meaning there is likely no one particular adaptative solution which suits all possible scenarios (Nelson et al., 2007). Historically, early life stress has always been a feature of human life (think hunter-gatherers), therefore the ability to adapt to stress has always been evolutionarily advantageous (Ellis et al., 2022). In such adverse conditions as humans used to live, very few children made it to adulthood, meaning that adaptation was likely always advantageous even despite a potentially high cognitive cost. It is highly likely that this ability to be highly adaptive to stressful situations, especially early in life, has been accentuated through natural selection (Ellis et al., 2022).

Life history theory is a particular example of how natural selection has shaped our developmental response to adversity. This theory considers the trade-offs an individual must make to succeed in life effectively. Trade-offs between domains such as physical and cognitive growth, reproduction and parenting, and maintenance of the body are all considered, with investment in each domain involving benefits and costs. Life history theory considers that previous experiences may influence the domains which an individual chooses to favour or disfavour. Two broadly different life strategies are fast versus slow. Individuals employing a fast life strategy may reach reproduction age relatively early in life and may produce a greater number of offspring meaning more variety of developmental outcomes, but at the cost of reduced health and longevity. By contrast, a slow life strategy may focus on greater investment in the parenting partnership with offspring production later and fewer. Crucial to this theory is that neither strategy is considered better than the other, instead they are both similarly adapted to differing contexts (Ellis et al., 2022). We can thus consider life history theory as an explanation for why environment may impact development and as an argument for adopting an adaptive approach to this relationship.

In addition to support from evolutionary theories, adaptive theory is also supported by work within the cognitive developmental field. Indeed, some researchers have reviewed various findings which support what they term developmental specialisation for specific environmental conditions (Ellis et al., 2017; Frankenhuis & de Weerth, 2013). Following the specialisation theory, Ellis et al. (2022) further outline the sensitisation hypothesis which assumes that cognitive adaptations developed by individuals from low SES backgrounds will manifest as advantages primarily in similarly low resource situations. This hypothesis may provide explanation for findings which show SES-related disadvantages in such children, since such studies typically take place in laboratories or other high-resource, low-stress settings. In addition, it provides a clear framework for hypotheses in future work hoping to investigate this further.

Whilst there are many questions still to be answered, some work which supports the adaptive framework already exists. For example, Belsky et al. (1996) found that 3-year-olds boys with a history of insecure attachment with a caregiver at 12 months remembered negative events (e.g. spilling a drink) more accurately than positive effects (e.g. receiving a gift), whilst the reverse was true for children with secure attachment history. The observed pattern of results might be believed adaptive when it is considered that enhanced recall of negative events might facilitate detection and avoidance of similar future events, which may occur more commonly for children with insecure attachments (Szepsenwol & Simpson, 2019). Participants in this study were all from low-risk backgrounds, indicating that adaptive patterns may be detectable even not at the extreme ends of risk. Whilst this supports an adaptive framework, a measure of attachment is not typically considered a feature of SES, however other work also supports adaptation in relation to SES features. Schliemann and Carraher (2002) found that young Brazilian street vendors, who were from low SES backgrounds characterised by undernourishment and very little education, could not complete arithmetic problems when presented in a school-like context but could quickly and successfully answer these whilst working as vendors. Interestingly, it was reported that children attempted to use written algorithms (as are learnt in school) when asked to answer questions in a school-like context, whereas they relied on mental strategies not learnt in school when working as a vendor. These findings thus support that children may adapt to their specific environment and suggests that context may be remarkably important for studies which investigate children's abilities.

Work from other domains also supports that children adapt in response to their experiences. Within the realm of executive functioning (EF), which is thought to be particularly associated to SES, findings suggest that EF abilities alter in relation to experience but may only become apparent when tested in specific contexts. In one study, adult participants' current context was manipulated to either elicit a high sense of environmental uncertainty or to remain controlled with no amplified environmental uncertainty, whilst a measure of childhood unpredictability was also gathered (Mittal et al., 2015). Results found that adults who had experienced more unpredictable childhoods showed greater shifting ability but worse inhibition than those who had had more predictable childhoods, though this was only apparent in conditions of uncertainty. This finding is consistent with the idea that more attention shifting may be advantageous in contexts where opportunities are fleeting, even if doing so has attentional costs (Ellis et al., 2017).

Other work has suggested that procedural learning is enhanced by experience of poverty, which may become apparent in contexts of high financial pressure, though working memory capacity may be reduced (Dang et al., 2016). In a context of poverty and exposure to violence, McCoy et al. (2015) found that 9- to 10-year-old children demonstrated faster though marginally less accurate performance on cognitive tasks, further indicating an adaptation towards more automatic processing. In addition to cognitive studies, there is evidence from animal and neuroimaging work which also supports an adaptive framework (see Ellis et al. (2017, 2020) for reviews).

One thing to note is that considering SES-related differences as adaptive versus deficit is a largely theoretical exercise. A deficit-based view relies on existence of a norm against which any differences are interpreted as dysfunction. In DCN research norms have typically been based upon Western children from relatively high socioeconomic background and the bias to view differences in development that vary from this norm as negative is likely rooted in a history of marginalisation and discrimination of certain communities (Garcini et al., 2022; Nketia et al., 2021).

Whilst a change in theoretical interpretation could be applied to much research in this area, some empirical approaches might additionally help support this theoretical stance. Research which considers performance across a range of tasks and various contexts might reveal patterns of findings which support that development in response to environmental factors is adaptive rather than disruptive. For example, if children who have grown up in a noisy household were found to perform better on a working memory task in noisy compared to quiet environments and the inverse pattern was found for children who live in quiet households, this may support that development is not worse for either group, but that is has adapted to be optimal in the context they are most likely to experience. Indeed, as Garcini et al. (2022) outline, it is crucial that DCN research considers that neurodevelopmental patterns may be ‘conditionally variable’ (p8); further emphasising the need to involve participants from a broad range of backgrounds so that relations can be fully understood. Other work which uses an individual-level approach to look at how profiles of relative strengths and differences may differ according to early experiences may be informative about how aspects of cognition are prioritised. For instance, if greater shifting abilities are developed at the cost of worse inhibition for individuals who had experienced more unpredictability, we would expect to see different profiles of relative shifting and lower inhibition abilities across individuals with experience of different levels of unpredictability. Research which found such patterns may not only be interpreted under an adaptive framework but would surely strengthen arguments for a reconsideration of the typical, deficit-based interpretation. The reduction of such theoretical biases is imperative in DCN research not only in the context of inclusion and social justice but also for improving the quality of the work itself, as biases can reduce the validity of findings (Garcini et al., 2022; Nketia et al., 2021).

In conclusion, an adaptive framework purports that SES-related differences may not be deficits but instead adaptations to different experiences which may be driven by a survivalist instinct. This suggests there may be no ‘typical’ phenotype across environments since all phenotypes are dependent on context (Ellis et al., 2022). An adaptive approach to understanding SES-related differences is supported both by evolutionary theories and empirical findings and could have important societal implications. Specifically, moving away from a deficit framing of children from low-SES backgrounds may impact the support children received throughout development and later in life. Future work which further investigates how neurocognitive strengths develop in response to different contexts could be particularly helpful in the development of interventions which work to further employ and enhance these abilities.

1.6.4 Work that has included diverse range of toddlers in research

Whilst there are many challenges which have led to gaps in existing literature investigating relations between early experiences and neural and attention development, some work has overcome these. This has largely been facilitated by advancement of data collection methods afforded by improved technology, but it also due to increased awareness of limitations and efforts of particular researchers and groups. This section will outline the key methods used in this thesis and outline where existing research has utilised these methods to address questions relating to the inclusion of a diverse range of toddlers in neurodevelopmental work.

1.6.4.1 Mobile EEG

Whilst EEG is a well-established neuroimaging method, mobile EEG systems are a relatively new development, with the number of systems available to researchers and clinicians rapidly increasing. Mobile systems are detached from fixed appliances, meaning such systems open doors for developmental research both inside and outside the laboratory (Mathewson et al., 2024). Inside the lab, mobile EEG systems enable participants more freedom of movement, which provides researchers with opportunities to apply novel approaches and ask new research questions about neural functioning (Troller-Renfree et al., 2021). Outside of the laboratory, mobile EEG systems also provide possibilities to conduct neuroimaging research in a wide range of settings, including the home (Troller-Renfree et al., 2021), schools (K. Xu et al., 2022), museums (Cruz-Garza et al., 2017) and even outdoors (Piñeyro Salvidegoitia et al., 2019). Given this flexibility, mobile EEG systems may also facilitate large scale data collection, which could lead to more informative and impactful research.

In addition, the advancements discussed may be helpful in increasing the diversity of neurodevelopmental research and in including groups for whom use of these methods is typically more challenging (Lau-Zhu et al., 2019). For instance, neuro-diverse individuals may have sensory, cognitive or motor sensitivities which mean data collection with traditional EEG methods is difficult, but whose impact may be attenuated by greater flexibility during data collection protocols (Lau-Zhu et al., 2019). The flexibility afforded by mobile EEG may also be particularly helpful in overcoming some practical difficulties involved with data collection from toddlers. Free movement enables young children to have breaks during an experiment without necessarily having to end EEG data collection as well as facilitating paradigms involving physical play and exploration. Both of these factors may lead to improved quantity and quality of data collection as they allow children to remain comfortable and engaged, in addition to enabling neural data collection during more naturalistic behaviours than traditional EEG systems. In relation to increased diversity, mobile EEG systems facilitate the use of neuroimaging in field settings, including in other countries or in lower socioeconomic status (SES) communities within the UK. Increasing representation within neuroimaging research would enable research to be more generalisable and may be important for future identification of sub-optimal brain development.

In fact, some research has already succeeded in using mobile EEG from developmental populations in low-income countries (Lockwood Estrin et al., 2022) and home environments (Troller-Renfree et al., 2021). Though there are now numerous mobile EEG systems, both these studies used an Enobio EEG system, which was selected for use in Troller-Renfree et al.'s study (2021) after rigorous trialling of various set-ups. This study also included a large sample (N > 400) of infants from low-income backgrounds, from which the authors developed a series of recommendations for EEG data collection and processing. These guidelines can be used in conjunction with other best practices to help researchers make decisions around electrode selection, reference configuration, frequency band choice, and data processing and analysis decisions (Bell & Cuevas, 2012; Troller-Renfree et al., 2021). Indeed, there are a small number of standardised processing pipelines now in circulation that have been developed for use with developmental EEG data; in their paper, Troller-Renfree et al. (2021) adapted the existing Maryland Analysis of Developmental EEG (MADE) pipeline (Debnath et al., 2020) for use with a low-density array of 20-channels. Whilst there are methods to handle artefact-laden data, it is crucial that the possibility for high-quality data is optimised during data collection, which can be facilitated by researcher skills to help settle the child and family (Hervé et al., 2022).

In development of guidelines, it is also important that psychometric properties of mobile EEG systems are checked, and some research has already investigated the feasibility, acceptability and reliability of using mobile EEG. Using a NeuroSky ThinkGear mobile EEG system, relative alpha, beta, delta and theta power measures were derived from a single electrode in children aged 10- to 17-years-old and in both young and older adults. While eyes were closed, reliability ranged from fair upwards (ICCs between 0.57-0.85), which increased to good when eyes were closed (ICCS 0.76-0.85) (Rogers et al., 2016). Comparison of three different EEG systems also revealed mostly moderate or good reliability for measures of absolute EEG power across systems in both a seated and walking condition (Oliveira et al., 2016). Notably, reliability was higher during the walking condition and for measures of alpha (ICCs 0.84-0.99) and beta power (ICCs 0.86-0.99) compared to theta (0.47-0.96) and gamma (ICCs 0.49-0.98). Crucially, both of these studies along with most previous work has been conducted with adults, yet due to unique challenges of data collection with young children, it is important that reliability is tested within these younger ages too. Indeed, some work has found good reliability for a variety of ERPs in infants and young children (Munsters et al., 2019; Webb et al., 2020), though only one have used a mobile EEG system. In one study, Haartsen et al. (2021) found mostly moderate reliability for N290 and P400 ERP amplitude and latency measures from young children aged between 30- to 48-months-old. In their large-scale study using the same Enobio mobile system, Troller-Renfree et al. (2021) found split-half reliability for 12-month-old infants was generally good across a range of absolute and relative power metrics, though alpha and theta power tended to have lower reliability than beta and gamma power measures. Exploration of how reliability differed in relation to the minimum number of 1-second segments of EEG included revealed that good reliability was achieved for absolute power measures with at least 20 trials, whilst for relative power only a minimum of 15 trials were required. Another study conducted in a clinical setting found that test-retest reliability of absolute frontal alpha power measures ranged from moderate to good (ICCs 0.40 – 0.71) in a sample of 70 10-year-old children using a 4-channel *Muse*™ EEG headband (InteraXon, Toronto, Canada), but relative measures were not considered (Xu et al., 2023). Such findings suggest that mobile EEG systems can produce reliable measures of neural functioning in neurodevelopmental research, though much more work in needed to understand how reliability differs across setting, age range, and experimental set-up.

In conclusion, mobile EEG holds great potential for including a more diverse range of participants in research across a wide variety of settings, thus may be particularly helpful in improving representation of neurodevelopmental toddler research. Troller-Renfree et al.'s (2021) project in particular has demonstrated how mobile EEG can be successfully used to collect high quality EEG data from a large sample of very young children. Nonetheless, there is still room to explore how similar can be applied with children of different ages (i.e. young infants, toddlers) and from different backgrounds, and to ascertain information about the most reliable methods and measures.

1.6.4.2 Novel eye-tracking methods

Eye-tracking is a method that is well-suited for measuring visual attention in young children, since eye-tracking tasks can be conducted without the need for complex verbal instruction and do not rely upon children's comprehension ability or motor skills (Karatekin, 2007; Richmond & Nelson, 2009). In this way, eye-tracking methods may provide more direct and objective measures of children's abilities than other behavioural methods (Karatekin, 2007; Sasson & Elison, 2012) and can be used in research with infants and young children. As eye-trackers are typically attached to the bottom of a screen which a child may passively, there are no requirements for the child to wear any equipment, nor the need to interact with anyone unfamiliar to them, enabling a broader community of infants to successfully participate in research (Karatekin, 2007).

Eye-trackers can also yield a huge amount of information about looking behaviour in infants which is richer and more detailed than older, traditional measures (Gredebäck et al., 2009). Because eye-trackers using technology to detect rapid eye movements, measures can be very precise and accurate, and may lead to better replication of findings. With eye-tracking methods, researchers are able to time-lock gaze behaviours to the appearance of stimuli or other events on a precise scale, and can determine the specific location of a gaze by examination of areas of interest (AOIs). As well as AOI-based measures that indicate where participants looked the most, other measures of looking behaviour, such as of peak look duration, mean look duration and reaction times, can be determined via use of eye-trackers. Programs such as Matlab and E-prime can be used in conjunction with eye-trackers to present task stimuli, collect data and process data, meaning eye-tracking methods can provide a huge amount of detailed information without making huge time-demands (Gredebäck et al., 2009). As completion of eye-tracking tasks do not tend to rely on language, motor or other abilities, this method also provides a tool which can be used at different ages, thus enabling investigation of how looking behaviours change over development. Given these characteristics of eye-tracking, such methods can facilitate greater understanding about visual attention processing assumed to be underpinned by looking behaviours.

In addition to these advantages afforded by eye-tracking, advancements in technology are allowing the development of novel methods which may further facilitate inclusion of a wider range of participants than traditional methods. One such advancement is the use of gaze-contingent tasks. Gaze-contingent stimulus presentation refers to tasks in which stimuli are designed to appear only if/ when gaze is detected in a particular area or time frame. For instance, a task might be designed such that the key stimuli (let’s say a face) only appears after the eye-tracker detects gaze in the centre of the screen. In this way, participants are able to somewhat control the pace of tasks, which may lead to less data attrition than tasks which only using timings for stimulus presentation. Data attrition could be improved due to both increased data availability from already participating participants and by reduced drop-out rates of other participants (Haartsen et al., 2021). As well as practical advantages, the use of gaze-contingent stimulus provides opportunities to ask new questions about how participants chose to direct and control their attention.

Though eye-tracking methods have provided a wealth of information about the development of visual attention, studies are somewhat limited in their capacity to ask broader questions about the structure of attention in toddlers. Most existing eye-tracking studies have compared group differences in performance on a single eye-tracking task using a small cohort, however tasks which tap a variety of skills are needed to investigate broader questions about different types of attention. Until recently, no studies or measurement batteries that tapped all three types of attention system outlined by Posner & Petersen (1990) were available for use in infants and toddler (as reported by de Jong et al., 2016). This prompted the development of the Utrecht Tasks for Attention in Toddlers using Eye Tracking (UTATE), which was designed to assess orienting, alerting and executive attention in toddlers (de Jong et al., 2016). Whilst this battery has been found to be reliable and to have some level of both construct and predictive validity (van Baar et al., 2020), robustness of tasks has not yet been examined. Additionally, sample sizes in this study were at times relatively small (n = 12 in the reliability sample) and the battery includes only three tasks. Thus, there remains work to be done to establish whether eye-tracking tasks can robustly measure effects when used in large-scale studies with toddlers.

1.6.4.3 App-based methods

Somewhat similar to advantages afforded by mobile EEG systems, app-based methods also provide opportunities for data collection on a large-scale and from a wide range of families, thereby holding great potential to improve diversity of developmental research. Apps are cheap to develop and can be used on mobile or tablet devices, meaning they are a relatively low-cost method of data collection. As most adults Worldwide have access to mobile devices which they use regularly (Phillips et al., 2024), app-based methods can also provide opportunities for more regular data collection, as parents or children may complete tasks, questionnaires or other app-based games as regularly as multiple times a day. In addition to data collection, apps may also be used as a tool for communication between researchers and families, and may thus facilitate resource and information sharing. This could further aid inclusion of families from a variety of backgrounds in research due to increased trust and improved relations. In an increasingly online-focussed World, app-based technologies hold potential to reduce some socioeconomic disparities associated with access to information around early child development (Crouse et al., 2023), via opportunities to increase parental knowledge and support, to provide learning opportunities for children , and to build understanding of child development across environments.

Indeed, the purpose of apps may be broadly divided into different sections. Some apps focus on supporting parents or improving parenting knowledge (Blakeslee et al., 2023.; Virani et al., 2021), some are focussed on gaining knowledge about children's development (Matsubara et al., 2022), and some aim to identify children who need support, whilst others aim to deliver this support (Griffith et al., 2020). Numerous apps have been developed for each of these purposes, though perhaps apps for use as a data collection tool for learning about child development remains under-utilised. In addition, many of the apps designed to include child-directed activities are focussed on children aged from two years upwards (Griffith et al., 2020), whereas there are far fewer for infants and younger children. Apps for children of these ages tend to instead be focussed on providing parent support (Virani et al., 2021), which leaves a gap within which app technologies are not yet being well-utilised. Even apps that do collect information about children's development tend to use a standardised, one-size-fits-all approach through which all children of a particular age are presented with the same activities to do. Whilst this is a simpler method for designing apps, this approach doesn't necessarily ascertain the more precise and accurate measure of children's abilities and may be more influenced by other factors such as sleep and temperament than more individualised approaches. There is also further work required for child development apps which are regularly used and helpful to families from a broad variety of backgrounds. Some apps have been developed for use across cultures and countries, such as 'Thrive by Five' which was designed to provide parents with culturally-relevant, science-based information about early child development and from whose development researchers have provided a framework for developing similar content (Crouse et al., 2023), though app-based tools are still developed predominantly for White, English-speaking women (Phillips et al., 2024).

In conclusion, great use is already being made of the advantages provided by app technologies in relation to child development, with apps focused on both providing parent support and improving children's learning. Within this work, however, there remain limitations which future work could aim to overcome, and a need to assess the usability, feasibility and acceptability of apps for all communities there are designed to serve (Crouse et al., 2023).

1.7 CONCLUSIONS

Despite the great potential for research which considers the relation between SES and early attention and neural development, there remain challenges with this area of research. Challenges include difficulties defining and measuring SES, collecting neurocognitive data reliability during toddlerhood, and involving families from a more diverse range of backgrounds in research.

1.8 OVERVIEW OF THE EXPERIMENTAL CHAPTERS

The overall aim of this thesis is to increase understanding of the association between early experiences and development during toddlerhood. Given the current limitations in this field, namely the lack of research which includes either toddlers or a diverse range of participants, this thesis has two key subgoals; these are 1) improving the diversity of participant samples in DCN research and 2) developing tools and methods for including a greater number of toddlers in DCN research. It does so in several ways, including both qualitative and quantitative, primary, and secondary data analysis, and utilising a range of different methods (including eye-tracking, EEG, focus groups, app data collection). Despite these various methods, the underlying motivations and focus remain the same: to develop tools and expertise for including a diverse range of toddlers in DCN research such that the relation between experiences and development can be better understood. In particular, this thesis focusses on the development of attentional processes as a domain which may be particularly related to early experiences, and uses a neurocognitive approach to build tools so that future research can investigate this further.

A few important points need to be noted. First, given that the research from this thesis took place during the COVID-19 pandemic, some studies and the overall thesis had to be adjusted in response to closures and restrictions of both research and childcare settings. Thus, where relevant, the following overviews for each of the chapters will also explain the impacts from the pandemic.

1.8.1 Goals of the thesis

1. Increase knowledge about associations between early experiences and development during toddlerhood.
    a. Improve methodologies for measuring toddler cognition.
        i. Extend existing knowledge of relations between SES and attentional processes in young children by assessing robustness of battery of eye-tracking tasks measuring visual attention in a large sample of young toddlers.
        ii. Focussing on neural markers which may be linked to executive attention, assess the test-retest reliability of a portable neuroimaging system for use with older toddlers.

iii. To facilitate filling the gap in toddler neuroimaging research, develop and evaluate the feasibility of a free-play neuroimaging design for use with toddlers.
iv. Develop a tool which can be used for remote data collection, which is developed with parental input to ensure relevance and user fit.

b. Improve diversity of DCN research
   i. Utilise community-involvement in the development of a tool for remote data collection with infants/ toddlers, developing meaningful relationships through which information can flow in both directions.
   ii. Use an adaptive strength-based, rather than deficit, approach to consider how the children's experiences may relate to development.
   iii. To facilitate future neuroimaging research in community settings, assess the reliability of a wearable, portable neuroimaging system which may be well-suited to use in different locations.
   iv. Develop a data collection paradigm and evaluate its success so that it may be used in community settings in future research, thereby reducing some barriers to research.

1.8.2 Chapter 2: Relation between socioeconomic status and profiles of visual attention in young toddlers

This chapter investigated the structure of visual attention in toddlers and how this may be related to their early experiences in life. Data in this chapter came from a large sample of 18-month-olds who watched a large battery of eye-tracking tasks designed to assess early visual attention. As this design had not previously been used with this age range, it was first established that it provided a feasible and robust measure of visual attention in a subsample of participants. In line with an adaptive framework of SES, a data-driven approach was then used to investigate relations between SES measures and children's profiles of visual attention.

1.8.3 Chapter 3: Tools in the real world: theta change reliability in older toddlers

This chapter assessed the test-retest reliability of several electroencephalography (EEG) measures in a sample of 3- to 4-year-olds using a portable neuroimaging system. There is evidence to suggest EEG power in the theta and alpha frequency range may be involved in processes related to executive attention and functioning, which is a domain that may be particularly related to SES. The test-retest reliability of these measures, however, has never (to the best of my knowledge) been assessed in toddlers/ young children. It is vital to assess psychometric properties of neurocognitive measures as these may have important implications for the potential impact of findings. In addition, psychometrics properties likely differ across contexts, therefore it is important to assess reliability across set-ups to inform about the suitability of using systems in different (i.e. community or field) settings.

1.8.4 Chapter 4: Tools in the real world: feasibility of a free play EEG design with toddlers

This chapter developed a paradigm designed specifically to involve a greater number of toddlers in neurocognitive research. A free-play neuroimaging data collection paradigm was evaluated to establish its success in collecting usable EEG data from toddlers using a portable EEG system. Given the lack of knowledge about brain functioning in toddlers, this chapter additionally considered task-related differences in theta and alpha power, adding to understanding about these measures during toddlerhood. Portable EEG systems may be used in community settings and a design which does not rely on a laboratory setup such as in this study may particularly facilitate such use. This may contribute to reducing barriers to research for some families and enable a more diverse range of toddlers to participate in neuroimaging research, which may enable future findings into this period of development.

This study was delayed and had to be significantly altered due to Covid-19 restrictions. The initial plan was to conduct data collection in community settings to assess the feasibility of conducting neuroimaging research in this way. Due to nation-wide closures of all settings, then significant rules and precautions about visitors, this project ultimately had to be moved into the lab, where I have attempted to mimic a preschool-type setting.

1.8.5 Chapter 5: App-based tool for remote data collection

This chapter investigated methods for increasing the diversity of participants in developmental research, through development of a scalable app-based measure of early development. Focus groups and questionnaire data were used to gain understanding about what factors parents consider important for research which uses an app-based tool, which could have influential implications for future development of this. Data collected via a current app-based tool were also analysed to assess the validity of this approach. This chapter is focussed overall on overcoming some of the practical barriers which may prevent people from certain communities from engaging in research and fits the with the overall thesis aims in two ways; (1) by developing a tool which may be used in large scale studies with toddlers and (2) by engaging families who may typically not be involved in research, thereby ultimately increasing the diversity and representation of participant samples in DCN research.

Whilst the data collection methods for data collection with the tool itself was online and could therefore be achieved remotely, the Covid-19 pandemic impacted the development and distribution of this app. Adding the app-based tool to distribution sites such as the App Store and Google Play Store was tricky and I experienced many delays with this, with minimal help available from usual IT support sources due to the increased demand (putting courses online, supporting people working remotely, etc) on them. Additionally, the original plan was to attend community settings from the onset of my PhD to build ongoing relationships through which the tool could be distributed, which is particularly important when trying to involve groups who do not typically engage in research (Garcini et al., 2022)*. Due to Covid-19 restrictions I was unable to do so for much of my PhD and instead could only take the first step towards community engagement in this research.*

1.9 DESCRIPTION OF DATASETS AND STATEMENT OF CONTRIBUTION

The current thesis includes data from two pre-existing datasets. The dataset analysed in chapter 2 was collected as part of the Developing Human Connectome Project (DHCP; http://www.developingconnectome.org/project/). Data were collected by a team of researchers based at St Thomas' Hospital in London, UK and associated with Kings College London, and were pre-processed by my supervisor, Professor Emily Jones. Structural equation modelling of this data was also done by Emily Jones; all other analyses were done by me. The dataset in chapter 3 was also collected prior to my PhD, though data collection was led by myself and a post-doctoral researcher in my lab group (Dr Teresa Del Bianco). Initial conversion of this data into FieldTrip format was done by another post-doctoral researcher, Dr Rianne Haartsen; all further processing and analyses were performed by me.

2 RELATION BETWEEN SOCIOECONOMIC STATUS AND PROFILES OF VISUAL ATTENTION IN YOUNG TODDLERS

A portion of this chapter has been published in the following article:

Braithwaite, E., Kyriakopoulou, V., Mason, L., Davidson, A., Tusor, N., Harper, N., Earl, M., Datoo-Partridge, S., Young, A., Chew, A., Falconer, S., Hajnal, J. V., Johnson, M. H., Nosarti, C., Edwards, A. D., & Jones, E. J. H. (2023). Objective assessment of visual attention in toddlerhood. *eLife*, *12*. https://doi.org/10.7554/eLife.87566

Abstract

Visual attention is an important mechanism through which children learn about their environment, and individual differences could substantially shape their later development. In addition, early experiences (often encapsulated as socioeconomic status (SES)) play a significant role in children's development. Considerable research has indicated a link between SES and executive functioning (EF), with early attention often considered a precursor of later EF; thus, there is clear motivation to investigate the association between SES and early attention. Indeed, a growing body of evidence supports that they may be related, with some suggestion that individuals' attention and EF abilities may adapt in response to their early experiences, though it remains unclear how this relation occurs in young toddlerhood. This study investigated how SES was related to profiles of visual attention in young toddlers, under an adaptive rather than deficit-based approach.

Participants in this study were 712 (333 females) 18-month-olds who were recruited as neonates to a large collaborative study (http://www.developingconnectome.org/). As this study utilised a large battery of eye-tracking tasks which had not previously been used with this age, it was first established that this method provided a feasible and robust measure of visual attention in the first 350 participants. Analyses showed expected condition effects for seven of eight tasks (*p*-values from <.001 to .04) and that quality and quantity of data collected was generally high. To build understanding about the underlying structure of visual attention in young toddlers and facilitate exploration of differences in the profiles of attention, structural equation modelling was applied across eye-tracking tasks. This indicated performance could be explained by two factors representing social and non-social attention, which is consistent with some theoretical models of visual attention.

Given that analyses of the first 350 participants indicated that this comprehensive eye-tracking battery was a feasible way of assessing visual attention in 18-month-olds, data from the larger cohort (n = 712) were then used to look at SES relations. A data-driven approach to investigating SES measures found two clusters which largely related to

typically low and high SES groups. Different patterns of performance were found in eye-tracking measures between these two SES groups. Specifically, latent scores relating to social and non-social attention did not differ for the high SES group, however the low SES group showed relatively lower social and higher non-social scores. Further investigation revealed that the higher SES group looked less to faces, were faster to find a hidden object in a working memory paradigm and were faster and more accurate in the learning phase of a cognitive control task.

Results supported that comprehensive eye-tracking batteries can objectively measure core components of visual attention in large-scale toddlerhood studies. Additionally, data-derived groupings can be used to investigate SES-related differences in performance on attentional measures, with evidence that profiles of visual attention differed in relation to SES experiences in early life. Findings are discussed within an adaptive framework.

Situate in thesis

The current chapter investigated relations between children's early experiences and development of visual attention abilities. It first established whether a large battery of eye-tracking tasks could be used to collect robust visual attention data from 18-month-olds, before exploring how SES was associated with profiles of visual attention at this age. A data-driven approach to SES grouping was utilised, in line with an adaptive rather than a deficit-based approach to the relation between early experiences and development. This fits with the aim of increasing diversity in developmental cognitive neuroscience research in my overall thesis, as a shift in the conceptualisation of experience-related differences in development may be important to reducing stigma for some communities. This chapter additionally contributed to the development of tools for use in large scale studies with toddlers and further provided information about how SES and visual attention are related in young toddlers.

2.1 INTRODUCTION

The first three years of life are critical for brain development (Fox et al., 2010) and have a significant impact on quality of later life and economic contributions (Knudsen et al., 2006). Optimising early brain development is thus a core goal of global public health policies (World Health Organization, 2022). Achieving this goal requires sensitive objective measures of infant cognitive development that can be used to examine the effects of different environments and policies on cognitive growth. However, most current approaches to large-scale measurement of early development are either through indirect and potentially imprecise measures such as head circumference, or through assessments of behaviour that are administered by skilled clinicians and may not be translatable to different contexts and cultures. This severely limits our capacity to track early cognitive development rapidly and objectively. An alternative approach is to leverage recent developments in technology that allow infant cognition to be assessed more directly. One highly promising method is eye-tracking, a non-invasive technology that is now widely used in small-scale laboratory studies of infant cognition. Eye-tracking can provide exquisite temporal and spatial resolution on an infant's direction of gaze and can be largely automated to produce scalable measures of individual differences in visual attention in infancy.

2.1.1 Visual attention and socioeconomic status

Visual attention is an important domain to measure because it provides insight into the mechanisms through which infants acquire knowledge from the world. Humans are a highly visual species and control information input through eye movements. Visual attention is a particularly crucial modality for learning in early development before infants have acquired advanced language or motor skills that provide alternative routes to learning. Whilst there is evidence that fundamental attention systems may be present at birth, these appear to alter and progress over development (see Colombo, 2001) such that it is possible that attention development may be impacted by early experiences.

Socioeconomic status (SES) is one way to conceptualise early experiences that is generally considered to be a measure of an individual's standing in society. SES incorporates a variety of social and economic factors, with (Antonoplis, 2023) defining it theoretically as "represent[ing] individuals' possession of normatively valued social and economic resources" (p279) and proposing that SES should be conceptualised as a combination of structural features which are each indicative of SES. Here, work which uses a range of measures relating to SES, such as household income, parental education, and parental occupation, was considered. Indeed, there is much existing work that indicates a relation between various measures of SES and cognition, with some findings in this area having identified a potential link between SES and visual attention which would benefit from further investigation. Whilst a small body of work has looked at associations between SES and visual attention directly, other support for this comes via a link between visual attention and executive functioning (EF). Considerable research has indicated that SES is associated with EF (Lawson et al., 2018; Raver et al., 2013), with growing evidence for this in infancy and toddlerhood. Given that EF and attention are closely linked in the first years of life, and early attention abilities are often considered precursors to later EF (Kraybill et al., 2019), it follows that there may be a relation between SES and attention.

Existing findings have pointed to a potential link between SES and several domains of visual attention including both social and non-social attention. In a paradigm which included a social condition, it was found that high SES (determined as maternal education involving at least one year of college education) infants showed more attention to people and toys during a free play task than low SES peers, with low SES infants showing more inattention (quiet disengagement or not looking) behaviours (Clearfield & Jedd, 2013). In addition, high SES infants showed greater increases in attention when stimuli became more complex. In work which considered more non-social aspects of attention, children from more advantaged families (determined based upon a composite measure of the primary caregiver's education, occupation and income) showed better accuracy attending to a visual cue and were more able to resist the interference of incongruent cues when compared to less advantaged children in an attention network task (Mezzacappa, 2004). Socially advantaged children also showed greater improvement in reaction time and accuracy in response to alerting cues. Various work also indicates a potential relation between SES and children's executive attention and executive functioning skills. At ages six, nine and twelve months, high SES infants (determined as in Clearfield & Jedd (2013)) showed expected developmental improvements in cognitive flexibility, whereas low SES infants showed a delayed response (Clearfield & Niman, 2012). Noble and colleagues (2007) further found that SES (measured via an income-to-needs ratio based upon education, occupation, and income) explained a significant portion of the variance in visuospatial skills, along with other neurocognitive abilities such as working memory and cognitive control. In a sample of nine-month-olds, a relation was found between SES (income-to-needs ratio based on education, occupation and income) and working memory skills, though this only appeared when objects were encoded with orienting, and not selective attention, processes (Markant et al., 2016). The authors suggested this may indicate that selective attention could mitigate the impact of SES on memory and may offer a potential mechanism for reducing its impact, thereby further indicating the potential use and importance of investigating SES-attention relations.

Evidence from neural work also supports an association between SES and attention. Selective auditory attention appears to be impacted by SES, with differences in ERP response to attended versus unattended stimuli found for higher, not lower, SES children at age four (Hampton Wray et al., 2017). In this work, families were specifically recruited from low and high SES backgrounds, where low SES families attended Head Start preschools and mothers of high SES children had a minimum education level involving college education. Age-related analyses further indicated that selective auditory attention was delayed and followed a divergent developmental pattern for the low SES compared to high SES group. Other neural work supports an impact of SES on auditory selective attention in cases where behavioural performance did not differ, with reduced ERP responses for low SES versus high SES groups when SES is based on various measures including a combination of caregiver education, family income and income-to-needs ratio (Kishiyama et al., 2009), a composite of parental occupation, education and measures of residential area quality (D'angiulli et al., 2012) and maternal education only (Stevens et al., 2009). SES (based on income-to-needs ratio) has also been associated with an ERP measure (the P3b component) relating to inhibition and attention allocation in 4.5- to 5.5-year-olds (St. John et al., 2019). In this study, higher P3b amplitudes were found to relate to higher income-to-needs ratio, but not maternal education, on both go and no-go trials of a go-no-go task. EEG oscillations have also been found to relate to socioeconomic measures, with lower frontal gamma power found for infants from lower SES (based on family income) compared to higher SES backgrounds (Tomalski et al., 2013). Whilst the functional role of gamma oscillations is not clearly delineated, it is thought to be involved in early language and attention development (Tomalski et al., 2013).

There is thus evidence of a relationship between SES and cognition, however most work has approached this from a deficit framework, whereby children from lower SES backgrounds are expected to perform worse than those from higher SES backgrounds. Whilst findings might provide evidence for this relationship, other studies show that it may be a more complex picture. Indeed, a growing body of research suggests that SES impacts on cognition may be better considered as adaptations to experiences, rather than necessarily deficits (Ellis et al., 2017).

2.1.1.1 Adaptive theory of SES and attention

Support for an adaptive theory of the relation between SES and attention comes from work involving both measures of attention and executive functioning (which may develop from early attentional skills).

Within the realm of EF, participants who had experienced greater unpredictability in childhood were found to have enhanced shifting abilities compared to those who had experienced more predictable childhoods (Mittal et al., 2015). Notably, this was only found under conditions of uncertainty and may be at the cost of poorer inhibitory control. It has been suggested that this trade-off between shifting and inhibition abilities may be particularly advantageous in unpredictable conditions where individuals may experience fewer and more fleeting opportunities (Ellis et al., 2017; Mittal et al., 2015) and may support that attentional adaptations, not necessarily deficits, are induced through experience of lower SES environments. It has also been suggested that working memory capacity may be diminished at the benefit of enhanced procedural learning in conditions of poverty (Dang et al., 2016), though findings are somewhat mixed regarding the relation between SES and working memory (Ellis et al., 2017). In a gap-overlap task designed to assess attention shifting abilities (Elsabbagh et al., 2009; Elsabbagh et al., 2013), five-month-old infants from higher SES backgrounds responded slower than lower SES infants in the overlap condition, where a peripheral stimulus competes with a central stimulus (Siqueiros Sanchez et al., 2021). Since faster responses are generally considered indicative of better processing, this finding contrasts with a deficit view of SES effects on attention. As the authors suggest, it may be that increased speed comes at the cost of accuracy, though further research is needed to investigate this further.

With these findings in mind, the current work takes a different approach to much SES-related work, utilising the large data set in this chapter to analyse whether attentional profiles may differ with SES. There is clear motivation to investigate how SES is related to attention in early development, as well as growing evidence that these may be considered adaptations to experience, rather than necessarily developmental deficits. This could be especially pertinent for the development of effective support for children who have had challenging early experiences.

Considering the work discussed thus far, there is clear motivation to further investigate a potential link between SES and visual attention. Indeed, both behavioural and neural work supports an SES-cognition association (Duncan et al., 1998; Hackman & Farah, 2009; Ursache & Noble, 2016), however less work has addressed this question using eye-tracking.

2.1.2 Eye-tracking as a measure of visual attention

As discussed in chapter 1 of this thesis, eye-tracking is particularly well-suited for measuring visual attention in young children, since eye-tracking tasks can be conducted without the need for complex verbal instruction and do not rely upon children’s comprehension ability or motor skills (Karatekin, 2007; Richmond & Nelson, 2009). Eye-tracking limits potential researcher bias as it minimises the role of the researcher and provides a more direct and objective measure of children’s processing (Karatekin, 2007; Sasson & Elison, 2012). Most eye-trackers used in research with young children are attached to the bottom of a screen which the child may passively watch whilst data is gathered. There are no requirements for the child to wear any equipment, nor the need to interact with anyone unfamiliar to them, enabling a broader community of infants to successfully participate (Karatekin, 2007).

Many developmental studies have utilised eye-tracking as a means of investigating attention. This has yielded important findings about how bias for faces changes over development (Leppänen, 2016), how orienting to stimuli differs under conditions of competition (Hood & Atkinson, 1993), and how memory guides attention (Hutchinson & Turk-Browne, 2012). However, such studies have typically focussed on comparing group differences in performance on a single eye-tracking task using a small infant cohort. Though insightful about the specific measures of these tasks, this approach cannot provide a wider understanding about visual attention. If we want to use eye-tracking to provide a comprehensive assessment of development on a large scale, we need to examine acquisition and task performance when we combine multiple tasks in an automated battery. Indeed, there are already some efforts to develop valid and reliable eye-tracking task batteries for use with toddlers (van Baar et al., 2020). Van Baar and colleagues (2020) found a battery of tasks designed to measure orienting, alerting and executive attention to be reliable and with some level of both construct and predictive validity, though they did not investigate robustness of tasks and were at times limited by a relatively small sample size (n = 12 in the reliability sample). Thus, there remains work to be done to establish whether eye-tracking tasks can robustly measure effects when used in large-scale studies with toddlers.

2.1.3 Structure of visual attention

Whilst the feasibility of using a comprehensive battery of visual attention tasks with toddlers requires investigation, using several tasks together in this way also provides advantages for assessment by allowing investigators to move beyond metrics extracted from single tasks. Taking a variety of measurements provides us with the opportunity to test whether the structure and profiles of attention across the battery fit with substructures of visual attention previously described in the existing literature. Much existing visual attention literature has separated environmentally-driven, bottom-up attention from internally-driven, top-down attention (Connor et al., 2004; Sarter et al., 2001). Commonly termed exogenous and endogenous attention respectively, it is thought they represent two independent attention systems which involve different behavioural patterns and which may be partially distinct at the neural level (Chica et al., 2013). Endogenous attention relates closely to executive attention systems that select goal-relevant actions by resolving conflict between competing inputs or impulses (Amso & Scerif, 2015; Miller & Cohen, 2001), whilst exogenous attention may be driven by properties of a stimulus such as luminance or complexity.

Whilst the distinction between exogenous and endogenous attention is relatively clear, conceptualisation of ‘social attention’ remains uncertain (Braithwaite et al., 2020). The term ‘social attention’ has taken on different meanings in different research papers, but generally refers to attention which focusses on a type of social stimuli, such as people or faces. A rich historical literature indicates that infants orient to faces from birth (Farroni et al., 2005) and show preference for face-like stimuli (versus non-face configurations), direct gaze (versus averted gaze or eyes-closed) and biological motion (versus inverted or scrambled) throughout development (Shultz et al., 2018). Indeed, some work has indicated the potential importance of early social attention in the emergence of neurodevelopmental conditions such as autism (Klin et al., 2015). However, the intersection between behaviours labelled as ‘social attention’ and the attention systems described above remains unclear (Braithwaite et al., 2020). Possibly, attention to social stimuli represents a specific case of attention to object features. Alternatively, social stimuli may gain attention due to the interplay between social motivation and other attentional systems previously mentioned (Chevallier et al., 2012; Dawson et al., 2005). Whilst such substructures have been examined longitudinally and through examination of differences between conditions, it remains unclear whether there are meaningful individual differences in separable subdomains of visual attention.

2.1.4 The current study

The current chapter focussed on investigating relations between SES and visual attention in young toddlers. It included a large cohort of 18-month-old infants tested on a large battery of eye-tracking tasks designed to measure both exogenous and endogenous aspects of attention in toddlerhood. To first establish that this combined delivery of eye-tracking tasks was a feasible way to collect visual attention data from this age of children, robustness was assessed in the data from the first 350 term-born toddlers. This was done by assessment of attrition rates, data quality and the presence or absence of expected condition effects. The structure of visual attention was also tested by fitting a theoretical model of visual attention to the pattern of data across tasks using a confirmatory factor analysis to determine if data reduction could be used and whether a smaller set of underlying constructs explained individual differences in visual attention within this cohort.

Having established robustness and visual attention structure in an initial sample of participants (the 'eye-tracking cohort'), further analyses used the resulting measures in combination with additional SES measures to explore associations between early experiences and visual attention in the whole sample of over 700 participants (the 'SES cohort'). To understand how different SES variables related to each other without reducing measures to a single composite score, in line with (Antonoplis, 2023) characterisation of SES as a combination of features related to the concept, a data-driven approach was used to group participants according to SES. Groupings were then used to explore the relation between SES and visual attention, which enabled investigation into whether particular profiles of SES were related to patterns of visual attention. SES-related differences in toddlers' visual attention skills were considered across tasks, to investigate whether eye-tracking profiles (i.e. mean looking time across different tasks) varied according to SES, as motivated by an adaptive, strength-based approach to SES.

Tasks in the eye-tracking battery were selected for inclusion based on previous literature and relevance to assessing broad types of visual attention in infancy. First, <u>exogenous attention</u> was measured through (i) the gap-overlap task (Elsabbagh et al., 2009; Elsabbagh et al., 2013), which measures the efficiency of shifts in attention from a central to a peripheral stimulus under competition and non-competition conditions; and (ii) a new non-social contingency task adapted from Kidd et al. (2012) in which the infant is presented with four balls, one of which will be activated (e.g. spin, make a sound) if the infant selects it with their gaze under different probabilistic reward structures; infants are expected to show faster saccades and greater attention engagement under moderately predictable versus random or completely predictable task conditions. <u>Endogenous attention</u> was measured through (i) the reversal learning task (or 'cognitive control', Wass et al., 2011) in which a brief video clip can be triggered if the infant selects one of two rectangles on the screen with their gaze; the lateral location of the video clip reverses after the first set of trials; (ii) a working memory task (based on the traditional behavioural working memory task in the Mullen Scales of Early Learning (Mullen, 1995)) in which an object appeared in the centre of the screen, then moved to one side and was covered by one of two curtains on either side of the screen; the infant could then 'find' the object by looking at the curtain behind which the object had disappeared; and (iii) a visual search task in which the infant had to 'find' a red apple amongst a set of 9 or 13 distractors of the same colour but not shape (complex) or neither the same colour nor shape (simple), adapted from Kaldy et al. (2011). Finally, social attention was measured through (i) the face pop-out task (Gliga et al., 2009), in which the infant views a series of slides containing a face, scrambled face, car, bird and phone while their gaze direction and duration is measured; (ii) the 'dancing ladies' social passive viewing task in which three women dance with objects (Saez de Urabain et al., 2017); in one condition the video is presented normally and in another it is phase-scrambled; (iii) the '50 faces' passive viewing task (Kamil, 2017) consisting of a series of brief clips of speaking male and female faces on a complex background.

Based upon work distinguishing between exogenous and endogenous attention (Connor et al., 2004; Sarter et al., 2001) and given uncertainty around how the construct of social attention relates to these (Braithwaite et al., 2020), a theoretical model of visual attention containing these three factors was fit to the pattern of data across tasks using structural equation modelling. This was compared to a simpler two-factor model which collapsed exogenous and endogenous attention into one non-social factor and included this along with a factor for social attention. After establishing robustness and comparing models in the eye-tracking cohort, it was then investigated how these metrics about structure and profiles of visual attention related to SES in the whole cohort, such that the association between toddlers' early experiences in life and visual attention could be considered.

2.2 METHODS

2.2.1 Participants

Parents were recruited from low-risk pregnancies as part of the developing human connectome project (*The Developing Human Connectome Project*, http://www.developingconnectome.org/). Children aged 18 months who had been recruited to the study as neonates were invited to attend an eye-tracking assessment and a neurodevelopmental assessment at the Centre for the Developing Brain in St Thomas' Hospital, London.

Two cohorts are referred to in this chapter. Exclusion criterion for both cohorts was the presence of a sibling with a diagnosis of autism spectrum disorder (ASD), whilst preterm infants (born at <37 weeks) were also excluded from the first ‘eye-tracking cohort’ so that feasibility could first be established in a typically developed, term-born sample.

The ‘eye-tracking Cohort’ therefore refers to the first 350 term-born children (184 males, 166 females) from whom eye-tracking data was collected. The second ‘SES Cohort’ included 712 children (379 males, 333 females) which included 566 term-born and 146 preterm-born participants. Cohort demographics are presented in Table 2.1.

The study was approved by the West London and GTAC Research ethics committee (REF: 14/LO/1169). Written informed consent was obtained from the parents/legal guardians of all participating children.

Table 2.1: Demographic data for two cohorts included in this chapter**.**

		Eye-tracking Cohort				SES Cohort			
		Mean	*SD*	Min.	Max.	*Mean*	*SD*	Min.	Max.
Gestational age at birth (weeks)		39.94	1.27	37.0	43.0	38.39	3.73	23.9	43.0
Age at assessment		18.23	1.13	17.0	24.0	18.99	2.31	17.1	34.2
		Count		*%*		*Count*		*%*	
Sex	Female	166		47.4		333		46.8	
	Male	184		52.6		379		53.2	
IMD quintile scores	1	25		7.1		72		10.1	
	2	32		9.1		78		11.0	
	3	59		16.9		128		18.0	
	4	113		32.3		212		29.8	
	5	97		27.7		169		23.7	
	Missing	24		6.9		52		7.3	
Maternal Ethnicity	Black	60		17.1		79		11.1	
	Asian	31		8.9		54		7.6	
	Chinese	3		0.9		24		3.4	
	White	217		62.0		458		64.3	
	Other	27		7.7		78		11.0	
	Missing	12		2.4		19		2.7	
Paternal Ethnicity	Black	56		16.0		94		13.2	
	Asian	31		8.9		57		8.0	
	Chinese	5		1.4		10		1.4	
	White	205		58.6		478		67.1	
	Other	40		11.4		52		7.3	
	Missing	13		3.7		21		2.9	
English first language	Both	161		46.0		393		55.2	
	Neither	97		27.7		175		24.6	
	Missing	24		6.9		19		2.67	
	At least one parent	68		19.4		125		17.6	

2.2.2 Materials

2.2.2.1 Eye-tracking

Raw eye tracking data was acquired from a Tobii Gaze Analytics SDK 3.0 (Tobii AB, Sweden) at a sampling rate of 120Hz, processed and saved to disk. Children sat approximately 60cm from the 23” screen (58.42cm x 28.6cm, 52.0° x 26.8°, native resolution of 1920 x 1080 pixels and an aspect ratio of 16:9) and stimuli were presented on Apple (Apple Inc., USA) MacBook Pro computers, using our custom-written stimulus presentation framework (Task Engine, sites.google.com/site/taskenginedoc/), running in MATLAB using Psychtoolbox 3 (Brainard, 1997; Kleiner et al., 2007) and the GStreamer library (gstreamer.freedesktop.org) for video decoding.

Trial onset and offset was associated with the current sample of gaze data, and time-stamped in the eye tracker’s time format. When a video was playing, an additional timestamp was recorded every 30 frames, to ensure constant synchronisation between stimuli and data. Eye-tracking variables are named to match the data release available via the Developing Human Connectome Project (DHCP; http://www.developingconnectome.org/project/).

2.2.2.2 SES measures

First language, ethnicity, highest education level and occupation information for both mothers and fathers were gathered via questionnaires. The Cognitively Stimulating Parenting Scale (CSPS) (Wolke et al., 2013) was also completed and index of multiple deprivation scores were derived from participant’s postcodes (National Statistics, 2019).

First Language: Parents were asked ‘is English your first language?’; where this was not the case, they were further asked ‘what is your first language?’.

Ethnicity: Parents were asked ‘how would you classify your ethnicity?’, with categories defined according to United Kingdom census categories.

Education: Parents were asked ‘at what age were you last in continuous education?’. Parental education is the age in years at which parents were last in continuous education.

Occupation: Parents were asked ‘what is your occupation?’, with a free-text box for responses. Parental employment statuses were transformed into broader job categories (such as teaching assistant, tutor, educator, pedagogical employee, translator >

educator), which were then transformed into occupation levels ranging from 1 (‘farm labourers’, ‘service workers’) to 9 (‘higher executives’, ‘major professionals’), according to Hollingshead's (2011) occupational scale.

Index of Multiple Deprivation (IMD) Score: This is a location-based composite SES measure derived from a participant’s address. Each postcode in the UK is assigned a score reflecting socioeconomic disadvantage in that area; postcodes from mothers’ addresses at the time of infant birth were used to find this score for each participant. A higher score reflects greater geographical deprivation.

Cognitively Stimulating Parenting Scale (CSPS): A questionnaire adapted from Wolke et al. (2013) designed to assess the availability of cognitively stimulating resources in the home. It includes questions about parental interactions, availability of educational toys and readiness of activities such as family outings, all of which promote cognitive stimulation. As participants were aged 18-months-old when this information was gathered, four items that are now commonly used when studying toddlers (i.e. relating to use of apps and tablets) (Bonthrone et al., 2021) were included. A total of 28 items were included in this version of the CSPS; scores from these can be aggregated to provide on overall score, referred to here as the 'home environment score’.

2.2.3 Procedure

All participants attended a visit at 18-months-old at which eye-tracking data were collected. SES data were either gathered during this visit or previously at enrolment into the study (see 2.2.4.1).

2.2.4 Eye-tracking

At the start of the eye tracking assessment the experimenter positioned each participant in front of the eye tracker. Online feedback was given to allow a position to be chosen as close as possible to the centre of the eye tracker head box, to maximise data quality. An automatic five-point calibration was then performed. At the beginning of each trial, a gaze-contingent fixation stimulus was presented in the centre of the screen; when gaze fell upon this stimulus the trial started. All tasks except for the Gap-Overlap began when the participant fixated a gaze-contingent central fixation stimulus, at a size of 3cm x 3cm (2.86° x 2.86° at 60cm viewing distance). If the participant became bored or fussy, the experimenter could skip the current trial and move on to the next. Skipped trials were marked in the data and excluded from analysis. Tasks were administered in blocks that were intermixed with each other and distributed throughout the battery; blocks of each task were presented interspersed to maintain child attention in a pseudorandomised order. Eye-tracking variables are named to match the data release, available via the Developing Human Connectome Project (DHCP; http://www.developingconnectome.org/project/).

2.2.4.1 SES measures

Sociodemographic measures were gathered at various timepoints; parental first language, ethnicity, education, and occupation information were gathered at parental enrolment into the study, IMD score was gathered at the birth of infants and CSPS score at the 18-month-old visit.

2.2.5 Stimuli

2.2.5.1 Eye-tracking

2.2.5.1.1 Gap-overlap

2.2.5.1.1.1 Stimulus presentation

Trials were presented in blocks of 12. Each trial started with the onset of a central stimulus (CS), a cartoon image of an analogue clock accompanied by an alerting sound. When the infant fixated the CS, after a 600-700ms wait period the peripheral stimulus (PS) was presented at 21.7° of visual angle on the left or right of the screen (random). In the baseline condition the CS disappeared concurrent with the appearance of the PS. In the overlap condition the CS continued to be presented for the duration of the rest of the trial. In the gap condition the CS was removed from the screen 200ms before the PS onset. The PS was a cartoon cloud that appeared on either the left or the right side of the screen and was accompanied by a sound, 3cm (2.86°) from the edge, rotating at 500° per second until fixated by the participant. A reward stimulus was then presented at the location of the PS for 1000ms. All stimuli were presented at a size of 3cm x 3cm (2.86° x 2.86° at 60cm viewing distance). Reward stimuli were either a star, sun, dog, cat, pig, tiger, or tortoise which were animated and accompanied by a sound.

2.2.5.1.1.2 Data extraction

Data were analysed offline. Each trial was inspected automatically to determine trial validity and calculate a saccadic reaction time (SRT) to shift attention from the CS to the PS, relative to PS onset. A trial was valid if the following conditions were met: 1) gaze fell on the CS; 2) no gaps of missing data longer than 200ms were present during the CS period (before PS onset); 3) there was at least one sample of gaze on the CS within 50ms either side of PS onset; 4) no gaps of missing data longer than 100ms were present during the PS period (between PS onset and reward onset); 5) SRT was longer than 150ms and shorter than 1200ms; 6) gaze did not go in the opposite direction to the side of the PS; 7) gaze did not enter the PS area of interest (AOI) after engagement with the CS but before PS onset.

Mean saccadic reaction times (SRTs; time of first sample to enter the PS – time of PS onset) were calculated for gap, overlap and baseline conditions (GO-Gap-SRT; GO-Baseline-SRT; GO-Overlap-SRT) separately, using only valid trials. For examination of condition differences, disengagement scores (GO-Disengagement) were computed as Overlap-SRT minus Baseline-SRT, and facilitation scores (GO-Facilitation) were computed as Baseline-SRT minus Gap-SRT. Reaction times are expected to be fastest in the gap then baseline then overlap condition (Elsabbagh et al., 2009).

2.2.5.1.2 Non-social contingency

2.2.5.1.2.1 Stimulus presentation

This task consisted of three blocks of 19 trials. There were three conditions: 0%, 60% and 100%. Blocks were distributed across the battery in a randomised order but constrained such that the 60% condition was always presented within the first two blocks. At the start and end of each block, a static picture of four balls was presented for a fixed period of 5 seconds. Between these was a set of trials. In each trial, participants were first presented with a fixation stimulus (hummingbird, 3cm x 3cm, 2.86° x 2.86°) that remained on screen until it was fixated by the infant. The fixation stimulus was replaced by a display of four balls (4.5cm, 2.3°), one in each corner of the screen, 3cm (2.9°) from each edge. Participants then saccaded to one of the four balls (within an AOI 50% bigger than the ball, subtending 4.3° degrees of visual angle); the ball they selected became their ‘chosen’ ball for that trial. The ‘reward’ for choosing a ball was that one of the balls would become animated; in the 100% reward block it was always the ball they chose that was animated; in the 0% reward block it was never the ball they chose (one of the other balls was randomly selected) and in the 60% reward block on 60% of the trials it was the ball they chose that was animated (and on the remaining 40%, a randomly chosen ball was animated). Rewards lasted 1000ms.

2.2.5.1.2.2 Data extraction

Offline: Saccadic reaction time to select a ball was computed as the difference between the frame on which the four balls appeared and the first sample to enter one of the four AOIs and averaged within a block. Fixation times to return to the central stimulus at the beginning of each trial were an index of engagement and computed as the difference between the frame on which the fixation stimulus appeared and the first sample to enter the fixation stimulus and averaged within a block. Trials were excluded if the reward was not played (to ensure the child had selected a ball) or if reaction time was less than 150ms (to avoid trials where children were looking at a ball when the trial started). Trials which had been skipped were also excluded.

Key dependent variables were saccadic reaction time to select a ball with removing reaction times > 2000ms ("zone outs"; NSC-100-SRT-NoZone, NSC-60-SRT-NoZone; NSC-0-SRT-NoZone); and fixation times to return to the central stimulus at the beginning of each trial (NSC-100-FixRT, NSC-60-FixRT; NSC-0-FixRT). Reaction times are expected to be fastest in the 60% condition because it is moderately predictable (Kidd et al., 2012).

2.2.5.1.3 Reversal learning ('Cognitive Control')

2.2.5.1.3.1 Stimulus presentation

This task was based on Wass et al. (2011). Two purple 17cm by 13cm (16.1° x 12.5° @ 60cm) rectangles were presented on the left and right of the screen (1.5cm, or 1.43° from the outermost edge). These remained on screen until either 1) either one of the rectangles was fixated by the participant, or 2) 2000ms had elapsed. At this point, one of the rectangles was replaced by video of the same dimensions, showing a 2s clip of the animated children's TV programme Thomas the Tank Engine. After the 2s clip had played, the screen became blank, and the next trial began with another gaze-contingent fixation stimulus.

In the first (non-scored) trial, the side that the child chose to look at first was recorded; if no side was chosen after 2000ms it was determined randomly. On the following set of 8 trials (termed the “learning phase”), the video was always presented on the opposite side to that chosen on the first trial. This learning phase ended after either a) the child made three anticipatory saccades to the correct side of the screen, or b) eight trials (excluding the first, non-scored, trial) had been presented. The task then entered the “reversal phase” where the correct side was reversed, and an additional nine trials were presented. The first trial of the reversal phase was not scored, but instead served to indicate (implicitly) to the child that the correct side had been reversed. Participants were presented with two blocks of 18 trials. Rectangles (17cm x 12.5cm, 16.1° x 11.9° @ 60cm) were 0.5cm (0.48°) from each edge of the screen and vertically centred.

2.2.5.1.3.2 Data extraction

Data were analysed offline. The first trial of each phase, where the participant was not yet aware of the location of the video, was discarded. AOIs were placed around the location of each of the rectangles (within one of which the video played) and dilated by 2° to account for poor calibration. Accuracy was calculated as *number of correct trials / total number of trials where an anticipation occurred* and saccadic reaction times were logged, then averaged across valid trials. Trials were considered invalid if SRT was less than 300ms or if no antisaccade was made.

Participants were excluded from analysis if they made fewer than 2 antisaccades (regardless of correctness) with valid SRTs per phase (learning/reversal), per block.

Key dependent variables were pre-switch accuracy (CC-Pre-Acc) and reaction time (CC-Pre-SRT). For assessment of task performance post-switch accuracy (CC-Post-Acc), and reaction time (and CC-Post-SRT) were also examined but because these were not completed by all participants and were dependent on performance in the pre-switch phase (missing not at random) they were not included in later modelling. Accuracy is expected to be lower and reaction times are expected to be slower during the pre-switch phase (in line with Wass et al., 2011). Pre-switch and post-switch anticipation (CC-Pre-Ant and CC-Post-Ant), calculated as *number of trials in which participants fixated a rectangle (whether it was correct or not)* / *total number of valid trials* within the pre-switch and post-switch phases, were included as control variables.

2.2.5.1.4 Working memory

2.2.5.1.4.1 Stimulus presentation

Images of two theatre stages with a lowered curtain were drawn on either side of the screen (16.0cm x 23.4cm, 15.2° x 22.1° @ 60cm). The stages remained presented through the trial. Over 500ms a toy appeared and dropped from the top to the vertical centre of the screen. Once the participant fixated the toy, the curtains on both the left and right stages lifted over 400ms. Over the next 750ms the toy moved to one of the (randomly chosen) stages. The toy remained motionless for 200ms before spinning for 400ms to engage attention at its stopping point on the stage. Over 400ms both stage curtains lowered, hiding the toy.

A central fixation was then presented until it was fixated by the participant; at this point it paused for 200ms before spinning for 200ms to engage attention and then disappeared. The task then waited for the participant to fixate one of the two curtains (defined as gaze being over the curtain for minimum 100ms); the choice was coded to be either correct or incorrect, depending upon whether the chosen curtain was hiding the toy. Over the next 400ms the chosen curtain was raised, revealing either the toy or an empty stage, depending upon whether the participant chose the correct side. If correct, the toy spun for 400ms as a reward before dropping off the bottom of the screen. The curtain then lowered. The reaction time (relative to the offset of the spinning central fixation stimulus) to choose a curtain, and the correctness of the choice, were recorded.

Participants were presented with three blocks of five trials. Theatre stage images were 16cm x 11cm, 15.2° x 10.5° and were presented 3cm (2.9°) from the edge. Trials began with the appearance of a horizontally centred child's toy, selected randomly from 19 exemplars. The toy was 7cm (6.9°) wide, and the height was adjusted to match the aspect ratio of each individual image.

2.2.5.1.4.2 Data extraction

Accuracy was calculated as *number of correct trials / total number of valid trials*. Saccadic reaction times were computed as the difference between the timestamps of the presentation of the first two rectangles and the first sample to enter a rectangle AOI, averaged across valid trials. Trials were considered invalid and were discarded if 2000ms elapsed without either curtain being fixated. The key dependent variables were accuracy of selecting the correct curtain (WM-Acc) and saccadic reaction times overall (WM-SRT-All). Reaction times for each of correct and incorrect responses (WM-SRT-Acc and WM-SRT-Inacc) were also compared. Infants were expected to achieve above chance-level accuracy (Daehler et al., 1976; Hofstadter & Reznick, 1996).

2.2.5.1.5 Visual search

2.2.5.1.5.1 Stimulus presentation

Search displays consisted of three different items; red apple (target, 4.57cm x 4.57cm, 4.3° x 4.3°), blue apple (colour distractor, 4.57cm x 4.57cm, 4.3° x 4.3°) and a red slice of an apple (an elongated rectangle, cropped from the full apple image, shape distractor, 1.12cm x 6.67cm, 1.1° x 6.4°). These stimuli were used to produce two trial types, feature search and conjunction search trials.

Stimuli in feature search trials varied on only one dimension, either colour or shape (e.g. a red apple surrounded by blue apples, or a red full apple surrounded by red slices). The set size in these trials was always 9, with one target stimulus and 8 distractors. Conjunction search trials consisted of an equal number of both colour and shape distractors, and the set size was either 9 or 13. In total there were six possible configurations of trial, two feature search trials and four conjunction search trials.

Stimuli were arranged on screen algorithmically. All stimuli were required to be within a circular region located at the centre of screen with a diameter of 27.2cm (25.5° @ 60cm). The target stimulus was then positioned at a random point within this circle, ensuring that it was not within 6cm (5.7° @ 60cm) of the centre of the screen, where it would overlap with the central fixation stimulus presented at the start of each trial. Next, each distractor stimulus was placed at a random location within the circle, ensuring that no stimulus (target or distractor) overlapped the spatial location of any other stimulus.

Each trial began with an animated central stimulus unique to this task. The target stimulus (a red apple) “flew” into the screen from one edge (left, top, right, bottom – chosen randomly) over 800ms, ending at the centre of the screen. When the participant fixated the apple, it faded into the background colour of the screen over 750ms, and then the full array was drawn. Note that the location of the target amongst the distractors was not the central location of this attention-getter, indeed the target was never presented within 6cm of this location. The array was presented for 4000ms or until the participant fixated the target stimulus, after which the target span as a visual reward for 1300s until the trial ended.

To highlight the special status of the target through pop-out, the first three test trials consisted of single feature displays. To emphasise this further and to grab participant's attention and fixation, before each trial began, the target (a red apple) 'flew in' from the upper portion of the screen, stopped in the centre of the screen for one second then disappeared. Aside from these first three, displays within each test trial were mixed in blocks and presented in random order. In all trials, a sound effect accompanied each event visual event.

2.2.5.1.5.2 Data extraction

Offline: AOIs were drawn around the target and distractors. Accuracy was calculated as *the number of trials on which the infant fixated the target (apple) before the animation was automatically triggered / the number of trials administered*. Saccadic reaction time to find the target on each trial was the difference between the time at which the search slide was presented and the time of the first gaze sample in the target AOI for each condition. Trials were excluded if reaction time was less than 150ms or if the trial had been skipped.

The key dependent variables are accuracy (VS-S9-Acc, VS-C9-Acc, VS-C13-Acc; and the saccadic reaction time to find the target on each trial VS-S9-SRT, VS-C9-SRT, VS-C13-SRT). Reaction times are expected to be fastest in the single search, then conjunctive 9-item search, then the conjunctive 13-item search (in line with Gerhardstein & Rovee-Collier, 2002).

2.2.5.1.6 Face pop-out

2.2.5.1.6.1 Stimulus presentation

Infants were presented with a series of six annular visual arrays each composed of five objects in different locations on the screen (Gliga et al., 2009; Hendry et al, 2018). Each array contained: a face with direct gaze, a visual 'noise' image generated from the same face presented within the array by randomising the phase spectra of the face whilst keeping the amplitude and colour spectra constant to act as a control for the low-level visual properties of the face stimuli (Halit, Csibra, Volein & Johnson, 2004), a bird, a car, and a mobile phone. Each array was presented for 10 seconds and counter-balanced for the location of the face in the array. The stimulus array was presented full screen with adjustments for a proper aspect ratio, at 43.8cm x 28.6cm (39.0° x 26.8° @ 60cm).

2.2.5.1.6.2 Data extraction

Areas-of-interest (AOIs) masks were placed around each stimulus. Each AOI was scored by counting the number of samples of gaze data that fell on each AOI. Trials were marked as invalid if either a) the proportion of valid (non-missing) samples was less than 25%, or b) the duration of data was less than 5s. For each AOI, the proportion of samples within it was calculated by *number of samples in AOI / number of valid (non-missing) samples.* Contiguous runs of samples within an AOI were identified and the mean proportion looking time and peak look duration to each AOI were calculated across valid trials only.

Peak look to each AOI is the duration of the longest look to that AOI during each trial averaged across the number of valid trials. Key dependent variables were the proportion of trials on which the infants first look to the face divided by the number of trials with a valid initial look (reflective of social orienting, Pop-Face-First) and percent looking to faces (Pop-Face-Pct, reflective of general interest (Elsabbagh et al., 2013)). There were also two key comparative AOIs for examination of the success of the manipulation, selected because they are the best visual control for faces (scrambled face or 'Noise') or are an object of interest (Car). These were percentage looking (Pop-Car-Pct, Pop-Noise-Pct) and peak look to each AOI (reflective of sustained attention (Gui et al., 2020) Pop-Face-Peak, Pop-Car-Peak, Pop-Noise-Peak). Percentage looking and looking times are expected to be greater to faces than comparison stimuli (Gliga et al., 2009).

2.2.5.1.7 Dancing ladies

2.2.5.1.7.1 Stimulus presentation

Each trial consisted of the presentation of one of six videos (20-25s duration each, 25 fps, 1920x1080 resolution) of three women dancing with objects. Videos were designed to be semi-naturalistic. Three videos were played in their native format and three matched videos were visually-scrambled such that the social information was degraded.

2.2.5.1.7.2 Data extraction

Offline: AOIs were hand traced around the faces and objects on a frame-by-frame basis using Motion (Apple Inc, USA) software, and were approximately 14.5° x 33.5° @ 60cm. Samples were assigned to AOIs and interpolated across gaps of < 200ms preceded and succeeded by the same AOI. Percent attention to each AOI was computed as *number of samples in each AOI/ total number of valid samples for that video* and averaged across videos within each condition; peak look to each AOI was defined as the longest run of samples within one AOI during each video, averaged across videos within each condition. Data was excluded if there were fewer than 25% valid samples for each video. Key dependent variables were percent attention to faces, Dance-Soc-Face-Pct, Dance-Scr-Face-Pct) and peak look (Dance-Soc-Face-Peak, Dance-Scr-Face-Peak). Additional comparators were percent attention to the object (Dance-Soc-Object-Pct, Dance-Scr-Object-Pct); peak look to the object (Dance-Soc-Object-Peak, Dance-Scr-Object-Peak). Looking times are expected to be greater to faces than comparison stimuli (Frank et al., 2014; Gliga et al., 2009).

2.2.5.1.8 Fifty faces

2.2.5.1.8.1 Stimulus presentation

Originating from the “50 People, One Question” project (Krolak, 2011), infants watched a video comprised of street interviews in English with a number of people (41s, 1280px x 720px, 25fps). The soundtrack of the original video was removed and replaced with classical music in order not to introduce linguistic confounds.

2.2.5.1.8.1 Data extraction

Offline: AOIs were hand traced around the faces, bodies and background people on a frame-by-frame basis using Motion (Apple Inc, USA) software. Samples were assigned to AOIs and interpolated across gaps of <200ms preceded and succeeded by the same AOI. Percent attention to each AOI was computed as *number of samples in each AOI/total number of valid samples for the video*; peak look to each AOI was defined as the longest run of samples within one AOI during the video. Key dependent variables were percent attention to faces (50Face-Face-Pct); peak look to faces (50Face-Face-Peak). Data was excluded if there were fewer than 25% valid samples for the video. To examine whether social attention was greater than to other elements of the scene, looking to the background (50Face-Background-Pct, 50Face-Background-Peak) was also examined. Looking times are expected to be greater to faces than comparison stimuli (Frank et al., 2014; Gliga et al., 2009).

2.2.5.2 Data quality assessment

In addition to task-specific data quality metrics (duration of valid data extracted from free viewing tasks and number of trials available from trial-based tasks), two measures of the general quality of the eye-tracking data across the session were computed. Accuracy (the spatial displacement of recorded gaze from the point fixated) and precision (variability in consecutive samples on the same fixation point) were extracted as proxies of general eye-tracking quality across the session.

Accuracy and precision were calculated during the gaze-contingent fixation stimulus that preceded each trial. The AOI around the fixation stimulus was 1.75x larger than the stimulus itself. The trial would begin even under conditions of high accuracy drift. Because the fixation stimulus was always at a known location, and because the trial would not begin until that location was fixated, it is possible to use it to calculate the spatial error between the true gaze location and the gaze location reported by the eye tracker. The AOI around the fixation stimulus was 1.75x larger than the stimulus itself. The trial would begin even under conditions of high accuracy drift. Accuracy was calculated as the root-mean-square (RMS) of the Euclidean distance between the location of each gaze sample and the location of the fixation stimulus. Precision was calculated as the RMS of the Euclidean distance between each gaze sample and the centroid of all gaze samples.

2.2.6 Statistical analyses

A range of variables pertinent to the valid acquisition of data from each task were first examined. These included the percentage of children who provided valid data; whether key metrics significantly varied with data quality (operationalised as having a correlation coefficient >= .2 or <= -.2); and whether the expected pattern of condition differences was elicited by the tasks administered in combination. Expected condition differences were assessed using Analysis of Variance tests (ANOVAs) or one-tailed t-tests where appropriate and where there was a strong directional hypothesis. Expected condition effects were: (i) Gap-Overlap: longer reaction times for the Overlap than Baseline and

Baseline than Gap condition (Elsabbagh et al., 2009; Elsabbagh et al., 2013); (ii) Pop-out: more looking at the face and a greater proportion of trials on which the first look was to the face (Gliga et al., 2009); (iii) Reversal learning: greater accuracy on the learning than reversal trials; (iv) Working memory: detection of the correct location significantly more than chance; (v) Non-social contingency: more attention to the last static slide and faster reaction times to select a ball during the trials in the 60% vs 100% and 0% conditions; (v) Dancing ladies: longer looking to faces than objects or background; (vii) 50 faces: longer looking to faces than objects, and show stronger effects for the native than degraded conditions; (viii) Visual search: faster reaction times and greater accuracy during the conjunctive nine condition. For all analyses, where sphericity was not met a Greenhouse-Geisser was applied and where Levene's test was significant, equal variances were not assumed. Central limit theory and the large sample size in this study meant normality might be assumed, though normality checks were also done to check no major deviations from normal. All data presented in this study will be made available via the Developing Human Connectome Project (DHCP) open access data release (http://www.developingconnectome.org/project/). Analyses were repeated after children with too few valid trials were removed ('cut-off analyses') and additional analyses were conducted to assess the impact of task-specific data quality metrics on condition effects. A hypothesis-driven SEM modelling approach was used to examine the underlying structure of visual attention (using lavaan (Rosseel, 2012)), before investigating how SES and visual attention were related. To do this it was first necessary to find out how SES features grouped. A hierarchical cluster analysis, using Gower distance measure, was conducted across SES measures. This analysis was chosen as the most appropriate clustering method for the size and types of variables included. Cluster groupings from this were then used to assess the relationship between SES and (1) underlying structures of visual attention and (2) profiles of visual attention across tasks.

2.3 RESULTS

2.3.1 Robustness of the battery

Robustness of the battery was assessed in the eye-tracking cohort (n = 350). Overall, both data retention and data quality were good, indicating that the current eye-tracking battery was successfully implemented. Tasks towards the end of the battery were skipped most often, indicating that prioritisation of the battery order was important, and some relations to data quality suggest care must be taken to report these variables in future work.

2.3.1.1 Data retention

Retention of data was good (Table 2.2). A high proportion of children (at least 70%) reached and provided at least some valid data for each task in the battery. For all but the non-social contingency task, over 90% of those who reached the task provided enough data to be included in condition effect analyses and for all tasks fewer than 10% of these participants were excluded after cut-off criteria were applied. Tasks towards the end of the battery were typically reached by fewer participants, though a high percentage (over 90%) of the children who reached these tasks did provide usable data.

Table 2.2: Data retention number and rates for each task for the eye-tracking cohort. All percentages are based on children who had enough data for inclusion on the highest yield variable

Task	*N* and % of children who got to task in battery (includes any participant who provided any valid data at all)		*N* and % with enough data for whole-group condition effect analysis		*N* and % with enough data for cut-off analysis	
Gap-overlap	337	96.3%	332	94.9%	302	86.3%
Non-social contingency	333	95.1%	253	72.3%	253	72.3%
Reversal learning	323	92.3%	323	92.3%	312	89.1%
Working memory	332	94.9%	331	94.6%	311	88.9%
Visual search	343	98.0%	341	97.4%	318	90.9%
Face pop-out	256	73.1%	256	73.1%	246	70.3%
Dancing ladies	251	71.7%	249	71.1%	226	64.6%
Fifty faces	246	70/3%	246	70.3%	235	67.1%

2.3.1.2 Data quality

Data quality was generally good (Figure 2.1). Accuracy and precision were 1.7 (*SD* = 0.8) and 1.5 (*SD* = 0.5) degrees respectively, with AOIs typically ranging from 4.3-39 degrees, indicating that the quality of the eye-tracking data was generally sufficient for the task design. Table 2.3 shows summary statistics of accuracy and precision data for the sample of participants who provided data for any task and for whom data quality data could be extracted. Despite cleaning and validation procedures, accuracy and precision did associate significantly with individual differences in key variables across many tasks (Table 2.4).

Figure 2.1: Raincloud plots for (a) precision and (b) accuracy of eye-tracking in the whole sample. All raincloud plots in this chapter were based on Allen et al. (2021).

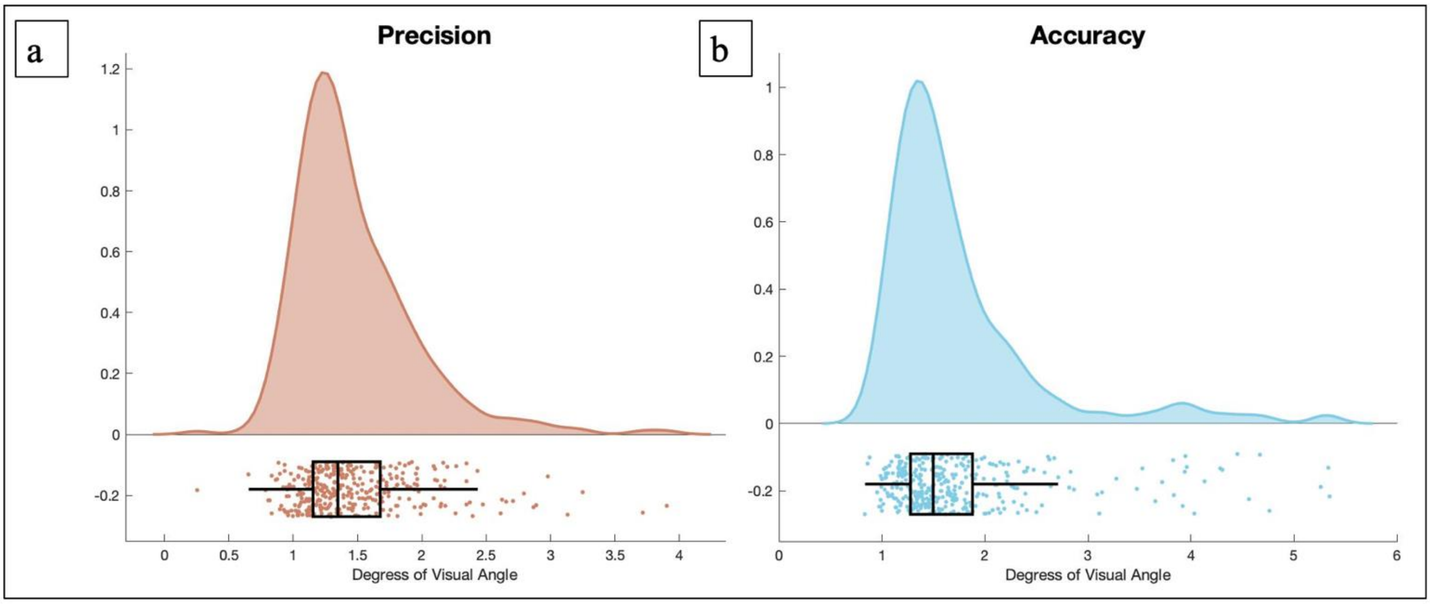


Table 2.3: Mean, standard deviation, range and number for precision and accuracy measures

	M	*SD*	Min.	Max.	*N*
Accuracy	1.74	0.79	0.83	5.35	340
Precision	1.47	0.47	0.25	3.90	340

Table 2.4: Associations between each of accuracy and precision, and key task variables; these associations include participants who provided enough data to be included in condition effect analyses. Cells are shaded according to r-values in the colour code below.

Task	Variable	Accuracy		Precision		*N*
		r	*p*	*r*	*p*	
Gap	GO-Gap-SRT	.25	<.001	.18	.001	333
	GO-Baseline-SRT	.11	.04	.02	.74	332
	GO-Overlap-SRT	.05	.36	.03	.61	333
	GO-Facilitation	-.15	.007	-.19	.001	331
	GO-Disengagement	-.01	.83	.01	.82	331
Non-social contingency	NSC-100-SRT-NoZone	-.06	.36	-.07	.27	282
	NSC-60-SRT-NoZone	-.09	.11	-.10	.08	323
	NSC-0-SRT-NoZone	-.25	<.001	-.20	.001	271
	NSC-100-FixRT	.26	<.001	.22	<.001	282
	NSC-60-FixRT	.22	<.001	.23	<.001	324
	NSC-0-FixRT	.19	.002	.27	<.001	271
Reversal	RL-Pre-Acc	.06	.32	.14	.01	320

Learning	RL-Post-Acc	-.05	.48	-.10	.15	203
	RL-Pre-SRT	-.05	.40	-.08	.16	320
	RL-Post-SRT	.03	.65	-.08	.27	203
Working Memory	WM-Acc	.05	.39	.11	.04	330
	WM-SRT-Acc	-.05	.33	-.13	.02	323
	WM-SRT-Inacc	-.05	.38	.13	.02	327
	WM-SRT	-.04	.52	-.15	.005	330
Visual Search	VS-S9-SRT	.06	.31	.01	.90	259
	VS-C9-SRT	-.07	.27	-.12	.05	257
	VS-C13-SRT	-.07	.30	-.11	.07	257
	VS-S9-Acc	-.32	<.001	-.16	.003	336
	VS-C9-Acc	-.15	.005	-.01	.86	336
	VS-C13-Acc	-.09	.12	.03	.61	336
Pop-out	Pop-Face-Pct	-.10	.12	-.10	.12	256
	Pop-Car-Pct	-.11	.09	-.06	.32	256
	Pop-Noise-Pct	-.09	.17	-.07	.30	256

	Pop-Face-Peak	-.12	.06	-.16	.01	255
	Pop-Car-Peak	-.14	.03	-.12	.05	254
	Pop-Noise-Peak	-.05	.46	-.11	.07	252
Dancing Ladies	Dance-Soc-Face-Pct	-.39	<.001	-.29	<.001	250
	Dance-Soc-Object-Pct	.05	.45	-.05	.45	250
	Dance-Scr-Face-Pct	-.13	.04	-.05	.44	250
	Dance-Scr-Object-Pct	-.17	.006	-.18	.004	250
	Dance-Soc-Face-Peak	-.37	<.001	-.30	<.001	249
	Dance-Soc-Object-Peak	.08	.19	-.07	.30	250
	Dance-Scr-Face-Peak	-.20	.001	-.17	.007	244
	Dance-Scr-Object-Peak	-.17	.009	-.22	.001	249
Fifty Faces	50Face-Face-Peak	-.21	.001	-.23	<.001	244
	50Face-Background-Peak	.05	.43	-.02	.78	228
	50Face-Face-Pct	-.28	<.001	-.19	.003	246
	50Face-Background-	.04	.50	.003	.96	246

Pct

Colour	Values
	0.1 <= |r-value| < 0.2
	0.2 <= |r-value| < 0.3
	0.3 <= |r-value| < 0.4

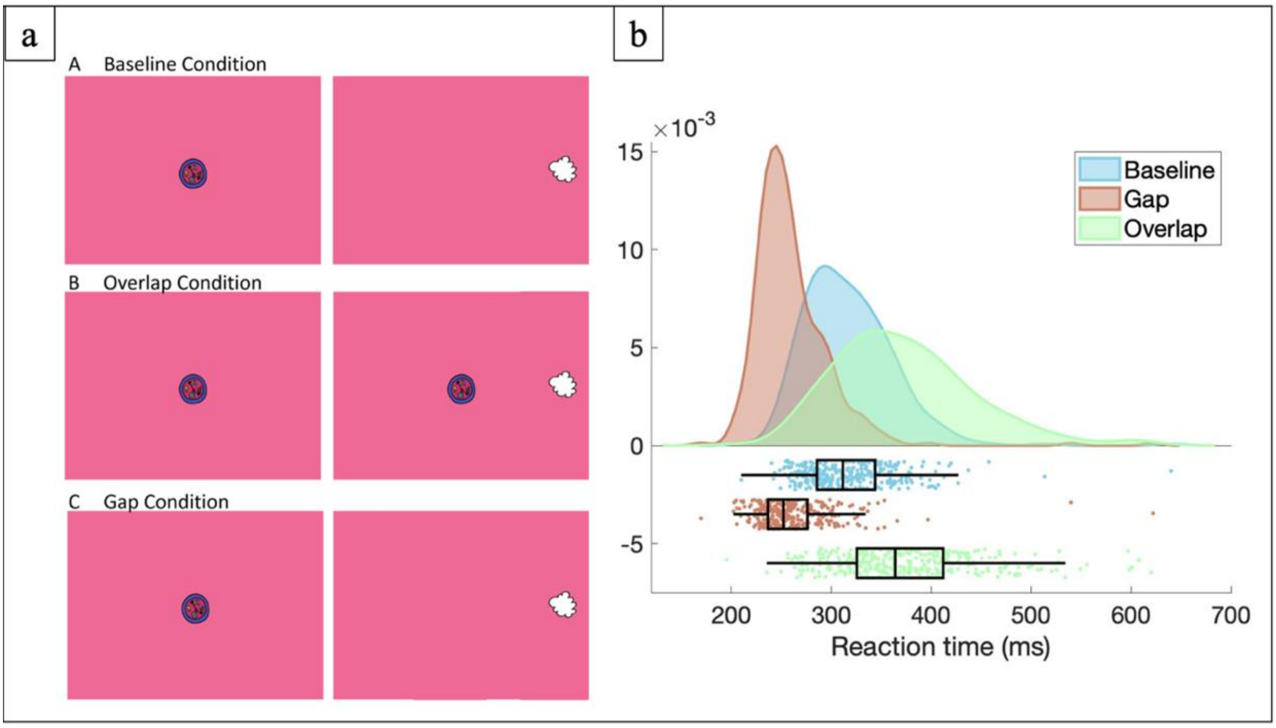
2.3.1.3 Condition differences

The expected pattern of condition differences was observed in each task, apart from the working memory task. Thus, each task (with the exception of working memory) replicated previously reported condition effects or showed the predicted pattern, confirming they could be robustly combined within a large-scale battery to measure differences in visual attention. Analyses for each task were repeated after exclusion criteria were applied ('cut off analyses'): exclusion criteria are listed in Appendix A. These showed that the same pattern of effects was found before and after exclusions were made, therefore analyses comparing visual attention and SES did not apply cut-offs (see Appendix A for analyses after exclusions). As a reminder, these analyses included data from the first 350 typically developing participants.

2.3.1.3.1 Gap-overlap task

94.9% of children successfully completed the gap task with an average of 13 valid trials per condition. A one-way repeated-measures ANOVA comparing reaction times in gap, overlap and baseline conditions indicated reaction times significantly varied with condition, ($F(1.68, 556.52) = 1015.76$, $n = 332$, $p < .001$, $\eta_p^2 = 0.75$); simple comparisons indicated (as expected) longer reaction times for the overlap ($M = 590.45$ms, $SD = 18.18$) than baseline ($M = 575.11$ms, $SD = 13.45$), ($F(1, 331) = 1056.474$, $p < .001$, $\eta_p^2 = 0.76$), and baseline than gap ($M = 555.30$ms, $SD = 13.33$), ($F(1, 331) = 372.501$, $p < .001$, $\eta_p^2 = 0.53$; Figure 2.2). Similar patterns were seen when restricting analysis to children with >5 trials/condition (Appendix A.1). Thus, this task yielded good data quantity and quality and the expected condition effects when administered in a longer battery.

Figure 2.2: (a) Gap-overlap task display and (b) raincloud plot of reaction times in gap,



overlap and baseline conditions of the gap-overlap task.

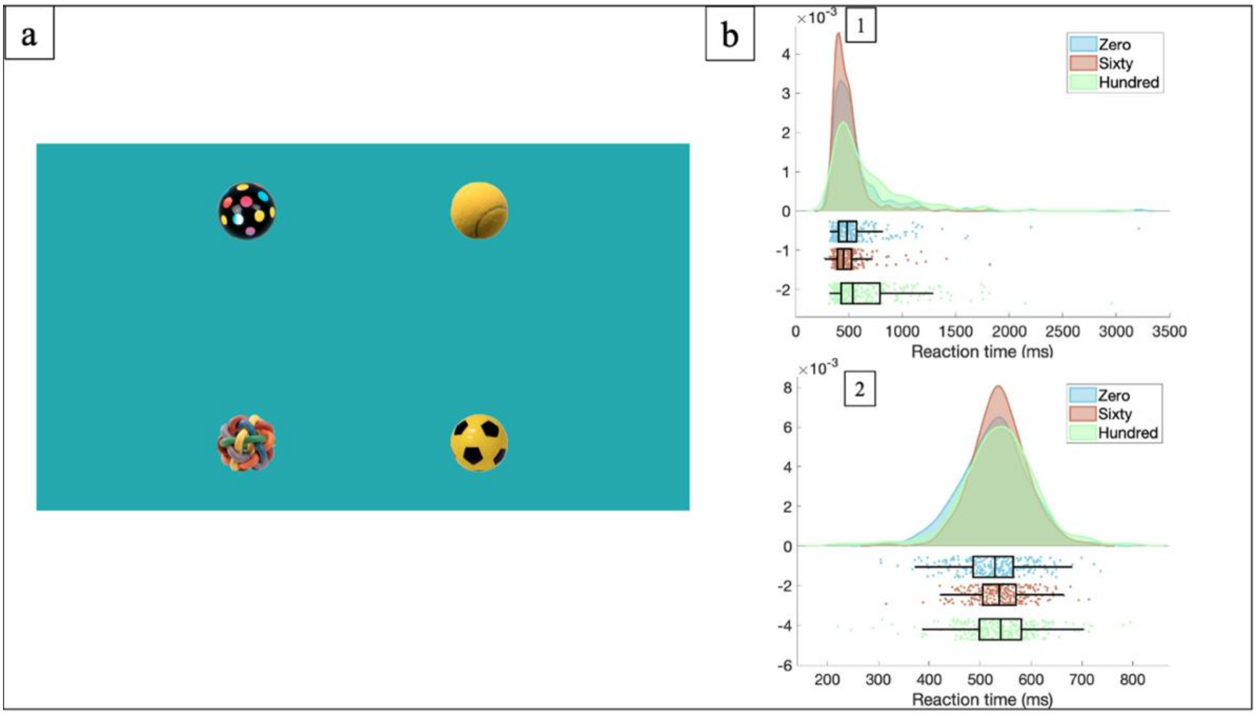
2.3.1.3.2 Non-social Contingency

72.3% of children successfully completed all three conditions with an average of 19 valid trials per condition. Multiple one-way repeated measures ANOVAs were conducted to compare two reaction time variables across three conditions (zero, sixty and hundred

Figure 2.3. Reaction time to select a ball did not differ between the sixty (*M* = 537.95ms, *SD* = 53.95) and hundred (*M* = 536.53ms, *SD* = 76.48), or between the zero and hundred conditions but, was significantly longer in the sixty versus zero condition as expected (*M* = 526.33ms, *SD* = 67.75); (Overall condition, $F(1.89, 473.55) = 2.93$, $p = .06$, $\eta_p^2 = 0.01$; sixty versus hundred, $F(1, 251) = 0.07$, $p = .80$, $\eta_p^2 = 0.00$; zero versus sixty, $F(2, 251) = 6.52$, $p = 0.01$, $\eta_p^2 = 0.03$; zero versus hundred, $F(1, 251) = 3.43$, $p = .07$, $\eta_p^2 = 0.01$). For times to return to the fixation stimulus, reaction times in the sixty condition (*M* = 488.06ms, *SD* = 170.78) were faster than in both the zero (*M* = 553.21ms, *SD* = 285.92) and hundred conditions (*M* = 661.82ms, *SD* = 352.26); (Overall condition, $F(1.89, 475.86) = 37.88$, $p < 0.001$, $\eta_p^2 = 0.13$; sixty versus zero, $F(1, 252) = 11.90$, $p = .001$, $\eta_p^2 = 0.05$; sixty versus hundred, $F(1, 252) = 59.69$, $p < .001$, $\eta_p^2 = 0.19$).

All children who completed the task met the cut-off criteria, therefore results were identical if children were only included with a minimum of 5 valid trials per condition (Appendix A.2). Thus, this task yielded reasonable quantity and good quality of data and expected condition effects were found when the task was administered in a longer battery.

Figure 2.3: (a) stimulus presentation and (b) raincloud plot of (b1) reaction time with zone-outs removed and (b2) reaction time to return to a central fixation stimulus after a



choice in zero, sixty and hundred conditions of the non-social contingency task

2.3.1.3.3 Reversal learning ('Cognitive Control')

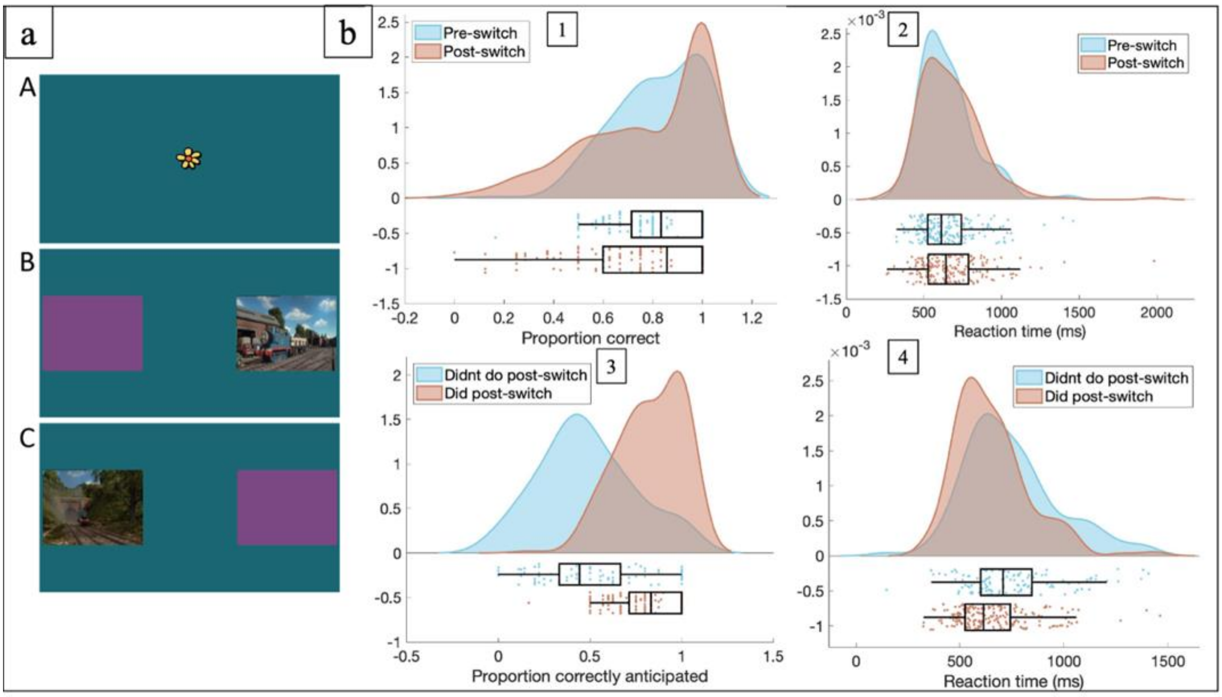
92.3% of children successfully completed the task with an average of six (*SD* = 2) valid trials for the pre-switch condition. During the pre-switch phase of the task, the mean proportion of trials in which participants correctly anticipated animation (CC-Pre-Acc) was approximately 70% (*M* = 0.70, *SD* = 0.27, *n* = 322). The mean reaction time for participants to select an AOI (CC-Pre-SRT) was 688.18ms (*SD* = 206.23, *n* = 322).

Of the group of 323 participants who completed the learning phase of the cognitive control task, 203 (62.8%) participants 'passed' and proceeded to the reversal condition, with an average of 5 (SD = 2) valid trials for this condition. Repeated-measures t-tests showed no significant differences in reaction times during the pre-switch (CC-Pre-SRT) (*M* = 653.70ms, *SD* = 183.42) and post-switch phases (CC-Post-SRT) (*M* = 670.61ms, *SD* = 205.53, *n* = 203), $t(202) = -1.08$, $p = .28$, CIs 95%[-0.05, 0.01], $d = 0.10$), however accuracy was significantly higher in the pre-switch phase (CC-Pre-Acc) (*M* = 0.83, *SD* = 0.17) compared to the post-switch phase (CC-Post-Acc) (*M* = 0.78, *SD* = 0.25, *n* = 203), $t(202) = 2.10$, $p = .04$, CIs 95%[0.003, 0.10], $d = 0.23$, Figure 2.4. Thus, children who completed both phases of the task showed some reduced performance after reversal, though not in all variables.

Independent t-tests conducted between the two groups who did and did not do the reversal phase found that as expected given the design, a significantly lower proportion of trials were correctly anticipated (CC-Pre-Acc) by the group who did not go on to do the reversal phase; this same group also showed significantly slower reaction times (Table 2.5 and Table 2.6).

Results were consistent if children were only included with a minimum of 2 valid trials per condition (Appendix A.3). Given not all children completed the reversal phase and data was not missing at random, further modelling thus used the accuracy and reaction time from the learning phase only. Overall, this task yielded good quantity and quality of data and some evidence to support the expected condition effects when administered in a longer battery.

Figure 2.4: (a) Reversal learning task display and (b) raincloud plot of (b1) the proportion of trials correctly anticipated and (b2) reaction times during the pre- and post-switch phases of the cognitive control task, (b3) the proportion of trials correctly anticipated and (b4) reaction time in the pre-switch phase by groups who did and didn't complete the post-switch phase.



Percentage of trials anticipated (either correctly or incorrectly) was included as a control variable. During the pre-switch phase, the mean proportion of trials in which participants anticipated an animation at all (CC-Pre-Ant) was 89% (n = 323, M = 0.89, SD = 0.19). A repeated t-test found that the proportion of trials anticipated in the pre-switch phase (CC-Pre-Ant); M = 0.96, SD = 0.10; was higher than in the post-switch phase (CC-Post-Ant); M = 0.93, SD = 0.16, $t(202)$ = 2.05, p = .04, d = 0.15, CIs 95% [0.0001, 0.05]. An independent t-test conducted between the two groups who did and didn't do the reversal phase found that a significantly lower proportion of trials were anticipated by the group who didn't go on to do the reversal phase, $t(142.25)$ = 7.28, p < .001, d = 1.01, CIs 95%[0.12, 0.22], means are in Table 2.6.

Table 2.5: t-value, degrees of freedom, p-value, Cohen's d, lower and upper CIs from independent t-tests between the proportion of trials correctly anticipated, anticipated at all and reaction times between two groups who did and didn't do the reversal phase in the cognitive control task

	t	*df*	*p*	*d*	95% CIs	
					lower	upper
CC-Pre-Acc	13.07	179.56	< 0.001	1.67	0.29	0.40
CC-Pre-SRT	-3.79	206.20	< 0.001	-2.34	44.71	141.90

Table 2.6: Mean, standard deviation and n for groups who did and didn't do the post-switch phase of the cognitive control task

Variable	Group	*M*	*SD*	*n*
Proportion of trials correctly anticipated (CC-Pre-SRT)	Did reversal condition	0.83	0.17	203
	Didn't do reversal condition	0.49	0.26	119
Reaction time (CC-Pre-SRT)	Did reversal condition	653.70	183.43	203
	Didn't do reversal condition	747.00	229.28	119
Proportion of trials anticipated (CC-Pre-Ant)	Did reversal condition	0.96	0.10	203
	Didn't do reversal condition	0.79	0.25	120

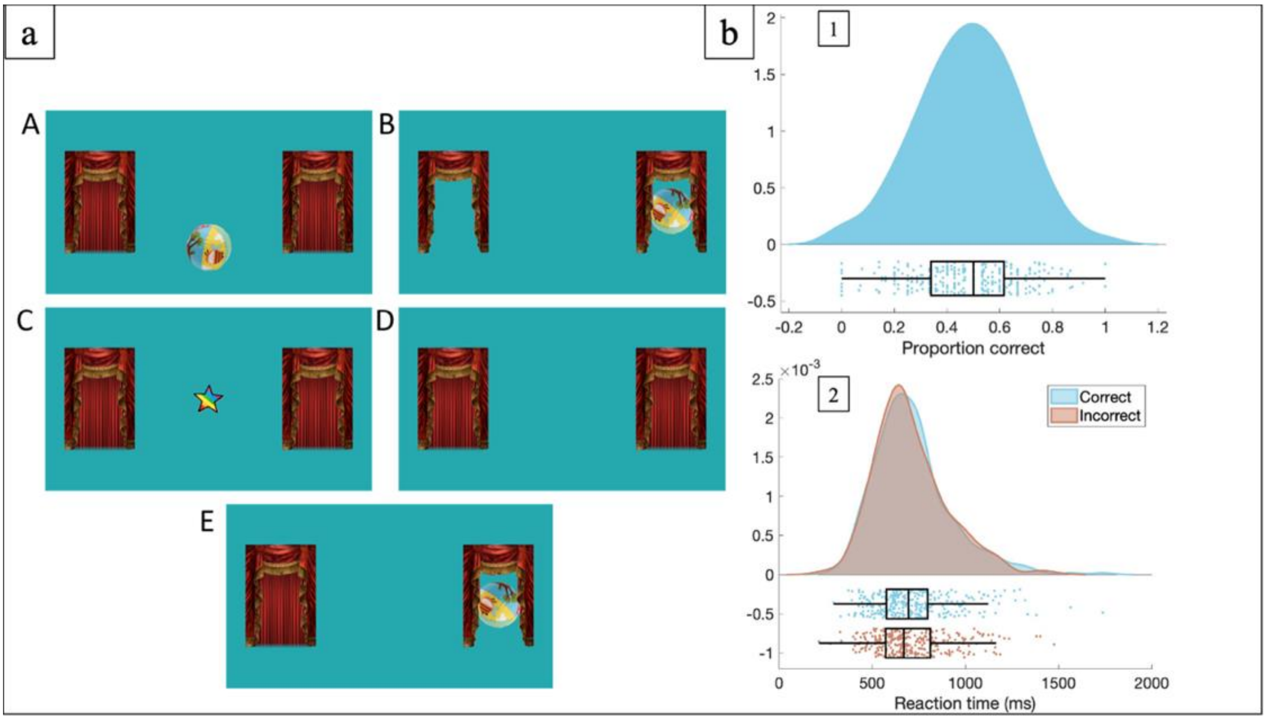
2.3.1.3.4 Working Memory

94.6% of children successfully completed the task with an average of 14 (3) valid trials. A one-sample two-tailed t-test revealed that the proportion of trials in which participants chose the correct location (WM-Acc) was not significantly different to chance, $t(330) = -1.91$, $p = .06$, [-0.04, 0.001], $d = -0.11$, $M = 0.48$, $SD = 0.19$.

A paired samples t-test showed that there was no difference in mean reaction times for trials in which participants were correct (WM-SRT-Acc) (*M* = 715.9ms, *SD* = 207.4, *n* = 321) versus incorrect (WM-SRT-Inacc) (*M* = 706.7ms, *SD* = 195.4, *n* = 321), $t(320) = 0.76$, $p = .45$, [-14.6, 33.13], $d = 0.05$, Figure 2.5). Mean reaction time across all trials (WM-SRT-All) was 700.8s (*SD* = 182.9, *n* = 324).

Results were consistent if only children with a minimum of 10 valid trials were included (Appendix A.4). Thus, children did not show evidence of successfully remembering the location of the object on the group level.

Figure 2.5: (a) Task figure and (b) raincloud plot of (b1) the proportion of correct trials and (b2) reaction times in correct and incorrect trials for the working memory task.



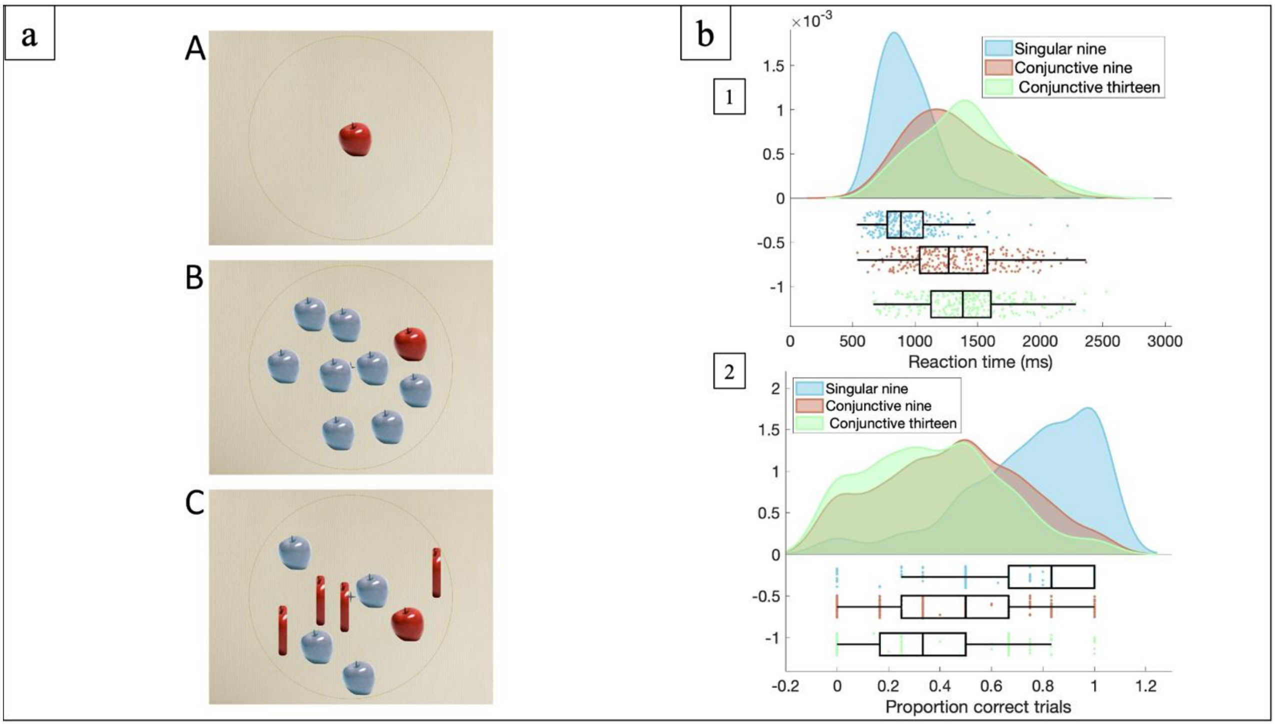
2.3.1.3.5 Visual Search

97.4% of children successfully completed the task with an average of 5 (*SD* = 1) valid trials per condition. Reaction times were slower in the conjunctive (*M* = 1316.68ms, *SD* = 36.12) than single feature displays (*M* = 938.79ms, *SD* = 241.21) with nine elements; (Overall effect of condition in ANOVA, $F(1.76, 450.90) = 159.90$, $p < .001$, $n = 257$, $\eta_p^2 = 0.38$; VS-S9-SRT versus VS-C9-SRT, $F(1, 256) = 255.46$, $p < .001$, $\eta_p^2 = 0.50$) and slower in the conjunctive thirteen (*M* = 1391.81ms, *SD* = 372.28) than conjunctive nine conditions; ($F(1, 256) = 5.62$, $p = .02$, $\eta_p^2 = 0.02$; Figure 2.6).

A similar repeated-measures ANOVA revealed there was also a difference in the proportion of correct trials across the three conditions: (F(2, 680) = 309.11, n = 341, p < .001, η_p^2 = 0.48). Pairwise comparisons revealed that the proportion of correct trials was higher in the singular nine condition (M = 0.75, SD = 0.26) versus the conjunctive nine conditions (M = 0.45, SD = 0.28); (F(1, 340) = 374.47, p < .001, η_p^2 = 0.52) and in the conjunctive nine versus the conjunctive thirteen conditions (M = 0.38, SD = 0.26); (F(1, 340) = 19.17, p < .001, η_p^2 = 0.05; Figure 2.6).

Results were consistent if children were only included with a minimum of 3 valid trials per condition (Appendix A.5). Thus, this task yielded good quantity and quality of data and found the expected condition effects when administered in a longer battery.

Figure 2.6: (a) Visual search task display and (b) raincloud plot of (b1) reaction times and (b2) accuracy to find the target in the singular nine, conjunctive nine and conjunctive thirteen conditions of the visual search task.



2.3.1.3.6 Face pop-Out Task

73.1% of children successfully completed the task with an average of 7 (SD = 2) valid trials. Analyses showed more looking to faces than cars and noise for both duration of looking; (Overall, F(1.39, 354.20) = 165.96, p < .001, η_p^2 = 0.39; face versus car, F(1, 255) = 60.31, p < .001, η_p^2 = 0.19; face versus noise, F(1, 255) = 607.29, p < .001, η_p^2 =

0.70) and peak look, (Overall, $F(1.43, 355.21) = 94.38$, $p < .001$, $\eta_p^2 = 0.28$; face versus car, $F(1, 248) = 8.43$, $p < 0.001$, $\eta_p^2 = 0.03$; face versus noise, $F(1, 248) = 363.16$, $p < .001$, $\eta_p^2 = 0.59$); all means in Table 2.7.

A one-sample t-test revealed that the mean proportion of looking time to faces was significantly higher than chance, $t(255) = 14.01$, $p < .001$, CIs 95%[0.10, 0.14], $d = 0.86$, where chance level was 1 in 5 (0.2) and mean proportion of looking to faces was 0.32, ($SD = 0.14$, Figure 2.7). The proportion of trials in which the first look was to the face was also significantly higher than chance, $t(255) = 26.50$, $p < .001$, CIs[0.37, 0.43], $d = 1.67$, M = 0.60, $SD = 0.24$, Figure 2.7. Results were consistent if children were only included with a minimum of 3 valid trials (Appendix A.6). Thus, the pop-out task had a reasonable quantity and quality of data and as expected showed higher face orienting and face attention than other comparative stimuli.

Figure 2.7: (a) Face pop-out task display and (b) raincloud plot of (b1) percentage looking and (b2) peak look duration to faces, car and noise, and (b3) proportion of trials in which the first look was to faces and (b4) percentage looking time to faces in the pop-out task.

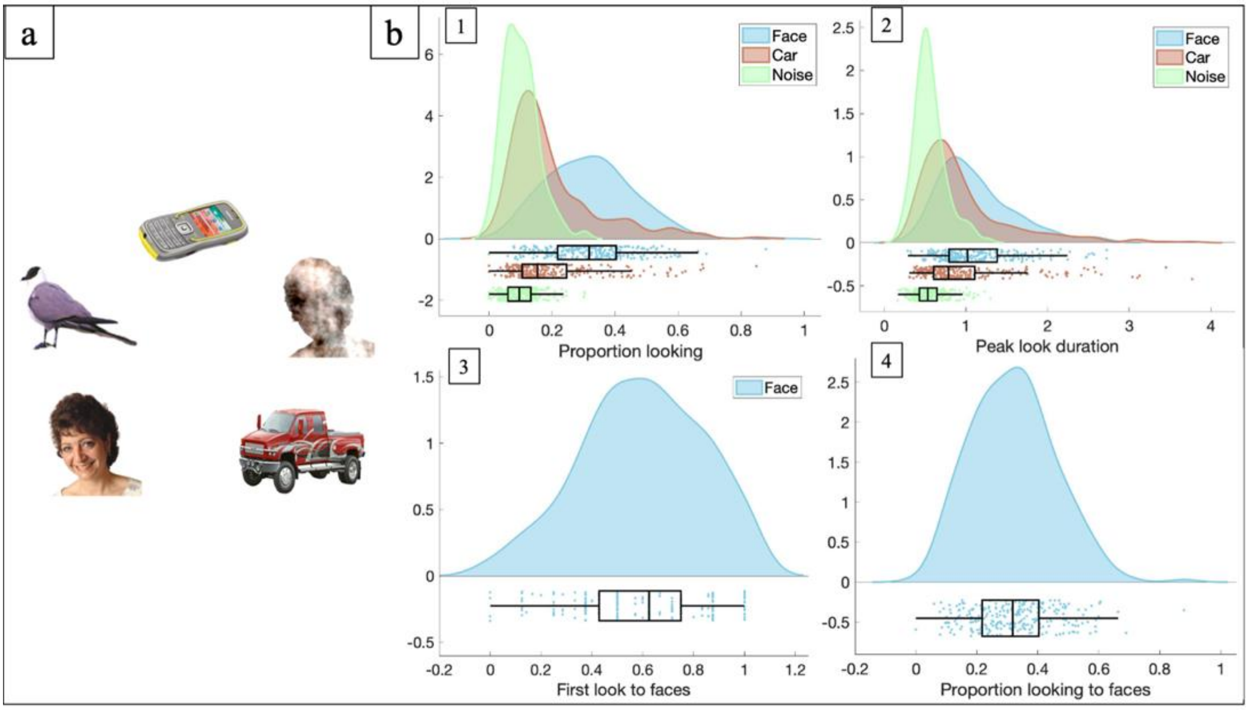


Table 2.7: Mean, standard deviation and n for overall looking and peak look duration to face, car, and noise AOIs in the pop-out task

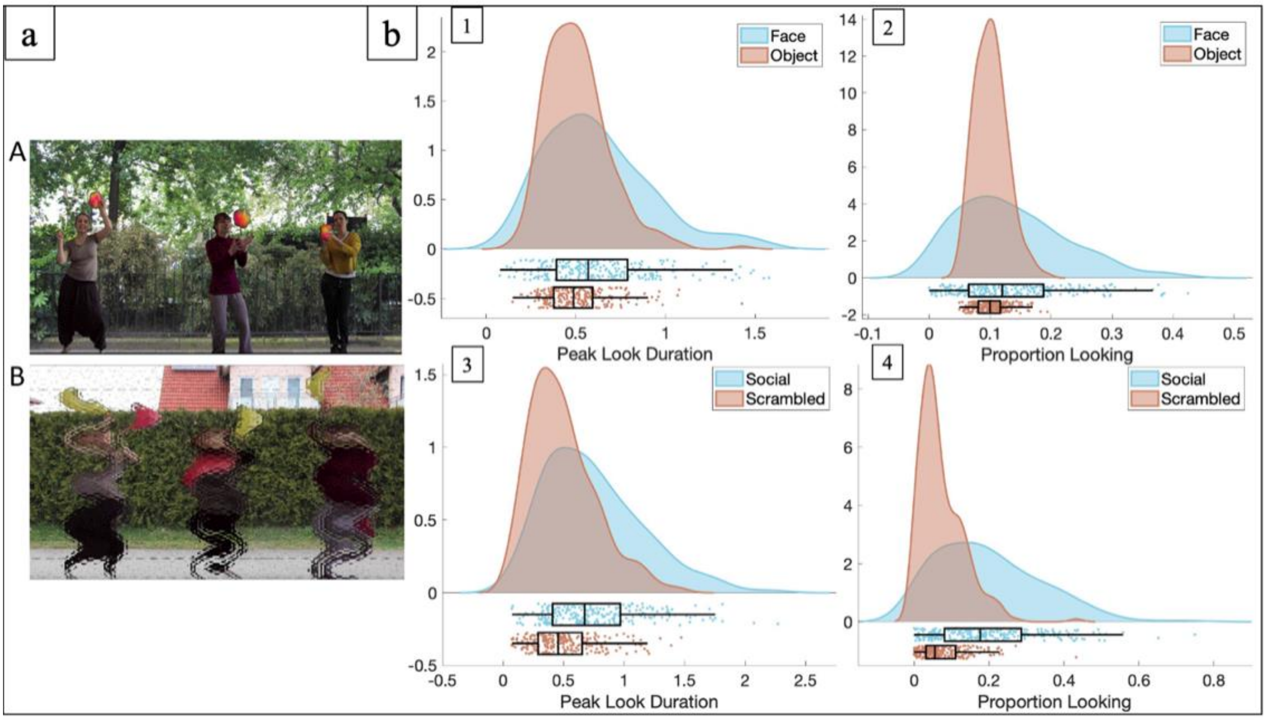
		Mean	*SD*	*N*
Overall looking	Face	0.32	0.14	256
	Car	0.20	0.14	256
	Noise	0.10	0.06	256
Peak look duration	Face	1.13	0.48	249
	Car	0.98	0.59	249
	Noise	0.56	0.19	249

2.3.1.3.7 Dancing ladies

71.1% of children successfully completed the task with an average of 4 (*SD* = 1) valid trials in each of the scrambled and social conditions.

A 2 x 2 repeated-measures ANOVA showed that peak look to faces (*M* = 0.62, *SE* = 0.02) was significantly higher than to objects (*M* = 0.51, *SE* = 0.01), (F(1, 241) = 24.93, p < .001, η_p^2 = 0.09), and peak look to faces was significantly greater in social (*M* = 0.75, *SD* = 0.41) than scrambled (*M* = 0.49, *SD* = 0.27), (*t*(242) = 11.36, *p* < .001, *d* = 0.75), Figure 2.8. Proportion looking to faces (*M* = 0.14, *SE* = 0.01) was also significantly higher than to objects (*M* = 0.10, *SE* = 0.002), (*F*(1, 248) = 29.66, *p* < .001, η_p^2 = 0.11), and proportion looking to faces was significantly higher in social (*M* = 0.20, SD = 0.14) than scrambled conditions (*M* = 0.07, *SD* = 0.06), (*t*(248)) = 16.67, *p* < .001, *d* = 1.21). Results were consistent if children were only included with a minimum of 3 valid trials of data (Appendix A.7). Thus, this task yielded reasonable quantity and good quality of data and found the expected condition effects when administered in a longer battery.

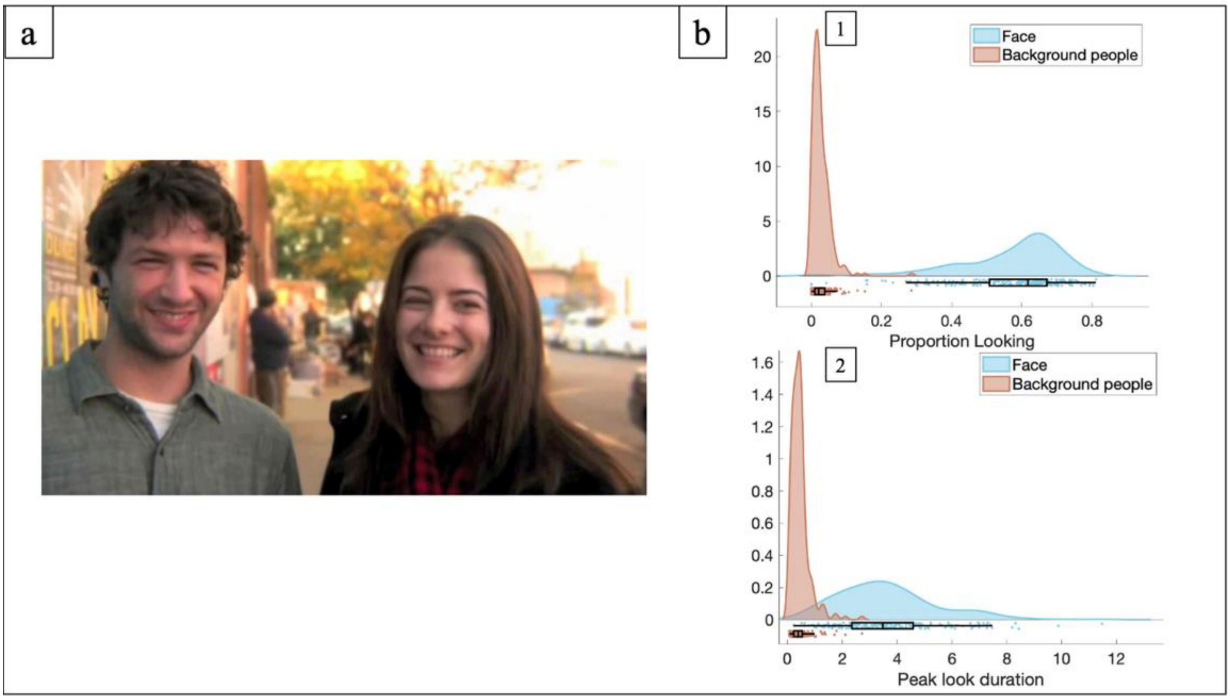
Figure 2.8: (a) Dancing ladies task display and (b) raincloud plot of (b1) peak look duration and (b2) proportion looking to faces and object across social and scrambled conditions, (b3) peak look duration and (b4) proportion looking time to faces in the social and scrambled conditions of the dancing ladies task.



2.3.1.3.8 Fifty faces

70.3% of children successfully completed the task with an average of 71.9 (25) percentage valid data. A paired samples t-test showed that peak look to faces (*M* = 3.65, *SD* = 1.83) was significantly higher than to background people (*M* = 0.46, *SD* = 0.34), (t(226) = 25.95, *p* < .001, *d* = 2.42). Proportion looking to faces (*M* = 0.57, *SD* = 0.15) was also significantly higher than to background people (*M* = 0.03, *SD* = 0.03), (*t*(245) = 53.83, p < .001, *d* = 4.99), Figure 2.9. Results were consistent if children were only included if they provided at least 80% of valid data (Appendix A.8). Thus, this task yielded reasonable quantity and quality of data and the expected condition effects when administered in a longer battery.

Figure 2.9: (a) Fifty faces task display and (b) raincloud plot of (b1) proportion of looking and (b2) peak look duration to faces and background people in the fifty faces task.



2.3.2 Factor Structure[1]

Bivariate correlations between core variables from each task (reaction time/accuracy in the cognitive control and working memory; reaction time in the gap; reaction time and proportion of trials retained in the non-social contingency; % looking and peak look to faces in the pop-out, dancing ladies and fifty faces tasks; accuracy in the visual search) were first examined. This showed strong intercorrelations between variables from the same task, and weaker correlations between tasks. Based on this information and literature-based hypotheses, a theoretically motivated structural equation model of the visual attention battery was conducted (using sem in Lavaan). A three-factor structure with three variables from each of three tasks per construct (proportion looking to faces, peak look and mean look duration to faces from face pop-out, dancing ladies and fifty faces), exogenous orienting (saccadic task reaction times in the three gap-overlap conditions, fixation reaction time in the three conditions of the non-social contingency

[1] Please note that these analyses were run by my supervisor, Emily Jones, not myself. Emily provided me with a social and non-social latent score per participant which I used in the analyses in section 2.3.3.1.

task, mu, sigma and tau parameters of the exGaussian modelling of reaction times) and endogenous attention (working memory accuracy and reaction time to correct and all trials, cognitive control accuracy in the learning conditions and reaction time during learning, and search accuracy across three conditions in the visual search; Figure 2.10a) was tested initially. This produced a good fit to the data; n = 194, RMSEA = 0.034, (CIs [0.019, 0.045], p = .99); CFI = 0.975; TLI = 0.9656; AIC = 11551; BIC = 12012; (c2 (294) = 358, p = .006, c^2/df = 1.21. However, model comparisons showed that a two-factor structure with the exogenous and endogenous factors combined did not fit significantly more poorly (c2 diff(2) = 1.61, p = .45) and was more parsimonious; thus the two-factor model, which contained social and non-social attention factors, was retained. Two data quality measures were also included based on a factor analytic decomposition of accuracy and precision and core data quality measures from each task; this revealed two factors reflecting accuracy and precision (21% of the variance) and median lost samples in the free viewing tasks (20% of the variance). The model produced a good fit using both multiple imputation to derive scores from all children (n = 350 RMSEA = 0.033 (CIs [0.025, 0.040], p = 1); CFI = 0.969; TLI = 0.958; AIC = 19340; BIC = 19988; c^2(296) = 408, p < .001, (c2/df=1.38) and complete cases only (n = 194, RMSEA = 0.034 (CIs [0.02, 0.046], p = .99); CFI = 0.974; TLI = 0.965; AIC = 11552; BIC = 12006; (c2 (296) = 363, p = .005, c^2/df=1.22; (Figure 2.10b). This model was a significant improvement on a model with only one latent factor which collapsed across social and non-social attention (n = 194, RMSEA = 0.037 (CIs [0.024, 0.048], p = .98); CFI = 0.970; TLI = 0.959; AIC = 11563; BIC = 12014; (c2 (297) = 376, p = .001, c^2/df=1.27; c2 diff(3)= 13.13, p = .0003). The two latent variables were not associated with each other (n = 194, B = 0.087, *SE* = 0.084, z = 1.04, p = .3). Higher scores for social attention represent more interest in social content; higher scores for non-social attention represent slower and less accurate cognitive responses (Figure 2.10b). These summary scores may prove useful for investigators wishing to represent distinct sources of variance in the battery as a whole.

Figure 2.10: Structual equation model of the structure of visual attention within the current eyetracking battery. Although our original model with three factors (a) provided a good fit to the data, it was not significantly better than a model with two latent variables

representing social and nonsocial attention (b)

2.3.3 Socioeconomic analyses

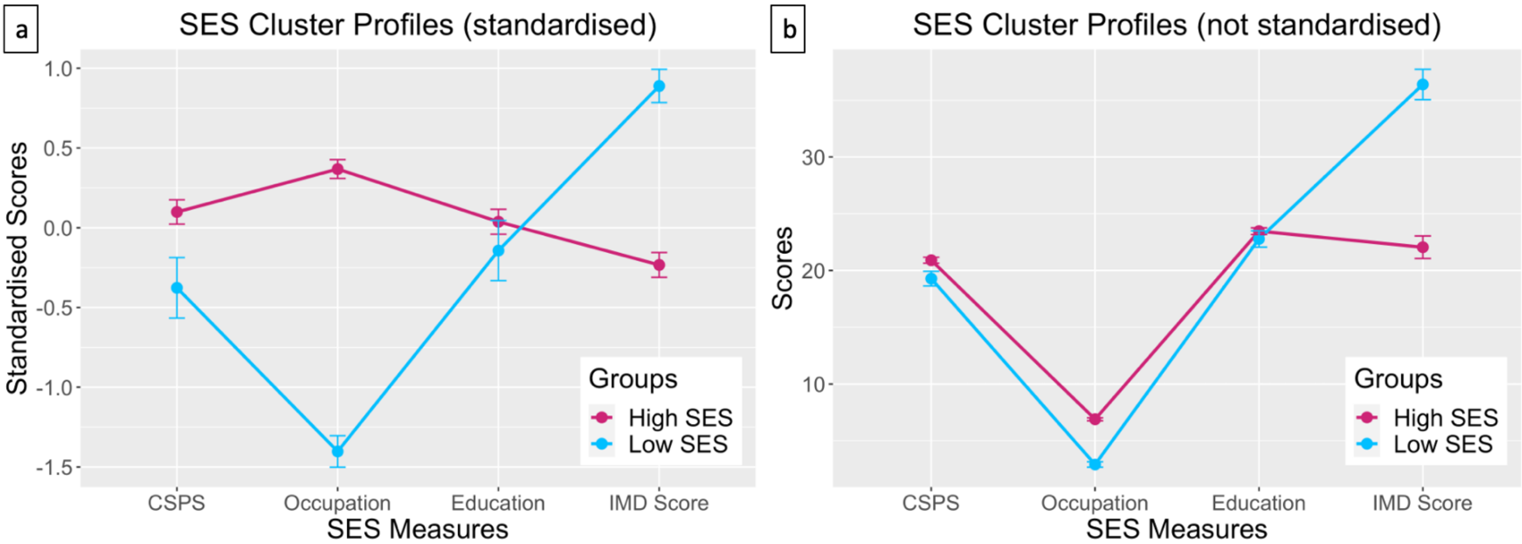
Any responses which were empty, ‘not given’, ‘NA’ or ‘-999’ were categorised as missing and missing data were then imputed using the ‘argImpute’ function in R (Harrell, 2023). Parental education and occupation measures were calculated as the mean of mother and father values and variables were scaled to have a mean of zero and a standard deviation of 1. A hierarchical cluster analysis was then conducted using the ‘hclust’ function in R (Nowakowski, 2023). Parental years of education, parental occupation level, IMD rank and CSPS scores were used; see Table 2.8 for further information about each of these metrics. The Gower distance metric was calculated using the ‘daisy’ function in R (Maechler et al., 2022) and a hierarchical cluster model was fit using the complete method. The elbow and silhouette methods were both considered to determine the optimal number of clusters; these indicated two clusters were optimal. Figure 2.11 shows how the scaled mean values of each SES measure differ across the two clusters. Cluster allocations were saved such that each participant was allocated to one of the two clusters.

Table 2.8: Mean, standard deviation and N for SES measures included in cluster analysis for cluster group one (high SES) and two (low SES)

	Group 1 (high SES)			Group 2 (low SES)		
	M	*SD*	*N*	*M*	*SD*	*N*
Cognitively Stimulating Parenting Scale	20.90	3.15	564	19.28	4.01	148
Occupation	6.89	1.58	564	2.91	1.47	148
Education	23.46	3.60	564	22.78	4.44	148
IMD Score	22.05	12.06	564	36.39	8.28	148

A Mann-Whitney test compared occupation scores between group 1 versus 2, whilst three separate independent t-tests compared CSPS, IMD and education scores. Levene’s test was significant for the three independent t-tests [$p < .001$], therefore equal variances were not assumed. CSPS and occupation scores for cluster one were higher than for cluster two: CSPS [$t(197.03) = 4.56$, $p < .001$, $d = 0.45$]; occupation scores [$U = 4146.00$, $z = -16.97$, $p < .001$, $d = 2.61$]. IMD score was lower for cluster two compared two cluster one [$t(329.61) = -16.88$, $p < .001$, $d = 1.39$], whilst there were not significant differences in education scores [$t(200.51) = 1.73$, $p = .08$, $d = 0.17$]. Given these differences and the SES profiles of each group, group 1 was named the high SES group and group 2 was the low SES group (though it should be noted that these are relative groups based upon data-driven groupings in the current sample and may not reflect high and low SES at the population level).

Figure 2.11: Mean scores for each SES variable in high SES (group one) and low SES (group two) for (a) standardised scores and (b) actual scores (error bars show 95% confidence intervals)



2.3.3.1 Data retention in SES Cohort

Retention of eye-tracking data in the SES cohort and across SES groups was also considered (Table 2.9). Data retention across groups was generally high and followed similar patterns, though attrition appeared slightly higher for the low SES group for some variables.

Table 2.9: SES Cohort: Number and percentage of participants who provided data for all variables included in SES analyses (2.3.3.3) for the whole SES cohort, high and low SES groups

Task	SES Cohort		High SES group		Low SES group	
	N	%	*N*	%	*N*	%
Gap-Overlap	663	93.1%	526	93.3%	137	92.6%
Non-Social contingency	499	70.1%	407	72.2%	92	62.2%
Reversal Learning	644	90.5%	514	91.1%	130	87.8%
Working Memory	669	94.0%	533	94.5%	136	91.9%
Visual Search	656	92.1%	521	92.4%	135	91.2%
Face Pop-out	534	75.0%	433	76.8%	101	68.2%
Dancing Ladies	524	73.6%	429	76.1%	95	64.2%
Fifty Faces	517	72.6%	419	74.3%	98	66.2%

2.3.3.2 Differences in the structure of visual attention across SES groups

A 2 x 2 repeated measures ANOVA was conducted to investigate whether underlying measures of the structure of visual attention differed according to SES cluster. Latent scores were entered as a repeated measures dependent variable with two levels (social and non-social), SES cluster group was an independent variable with two levels (high and low SES) and cohort was additionally entered as an independent variable with two levels (preterm and term).

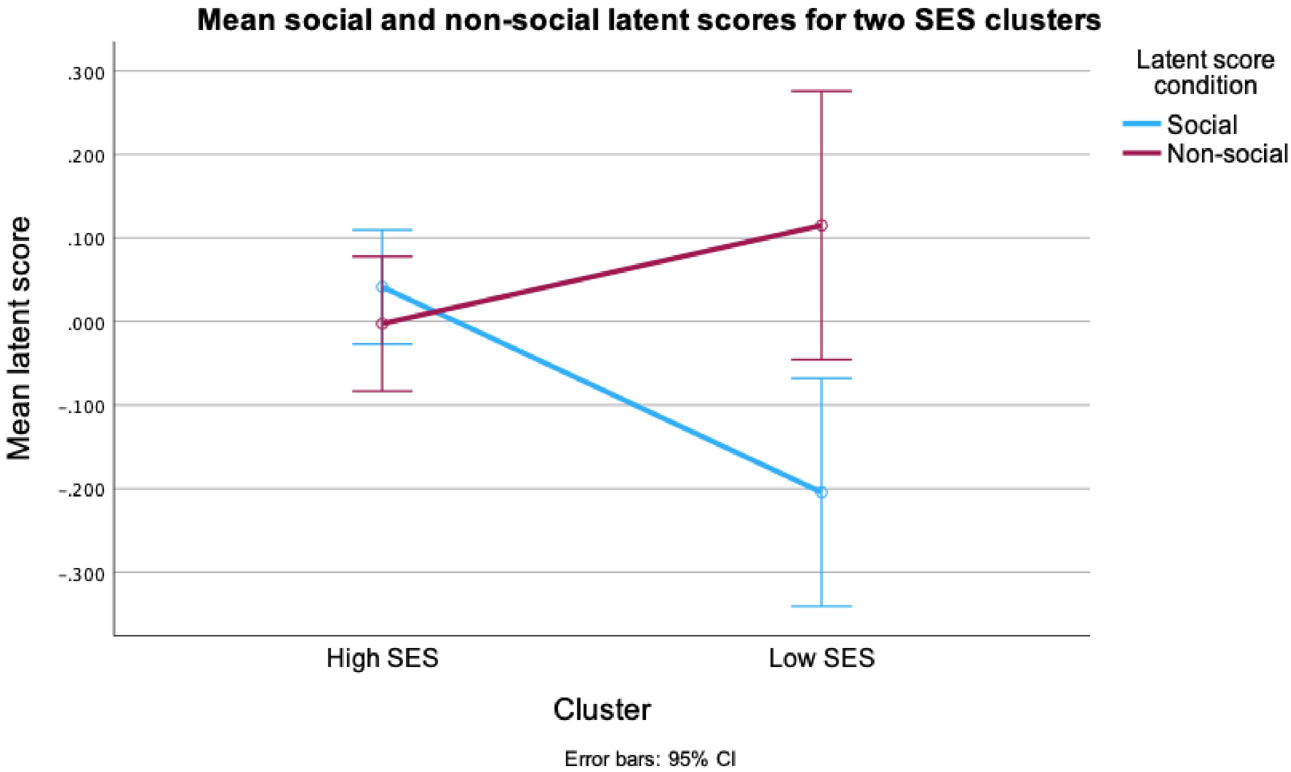
Cohort was included as an independent variable to assess whether preterm birth impacted the structure of visual attention; descriptive statistics of social and non-social scores for each SES group and the whole sample are in Table 2.10. The main effect of cohort was not found to be significant [$F(1, 708) = 0.13$, $p = .72$, $\eta_p^2 < 0.001$], nor were interactions between cohort and latent scores [$F(1, 708) = .45$, $p = 0.50$, $\eta_p^2 = 0.001$], between cohort and SES scores [$F(1, 708) = 2.25$, $p = .13$, $\eta_p^2 = 0.003$], or between cohort, SES and latent scores [$F(1, 708) = 0.01$, $p = .93$, $\eta_p^2 < 0.001$]. The main effect of latent score did not reach significance at the p < .05 level [$F(1, 708) = 3.86$, $p = .05$, $\eta_p^2 = 0.005$].

The main effect of SES was not significant [$F(1, 708) = 1.79$, $p = .18$, $\eta_p^2 = 0.003$], though there was a significant interaction between SES and latent scores [$F(1, 708) = 6.72$, $p = .01$, $\eta_p^2 = 0.009$]. Follow-up pairwise comparisons revealed no significant difference between social (*M* = 0.04, *SE* = 0.04) and non-social (*M* = -0.003 *SE* = 0.04) latent scores for the high SES group [*Mean diff* = 0.04, *p* = .48, CIs 95%[-0.08, -0.17] but that non-social latent scores (*M* = 0.12, *SE* = 0.08) were significantly higher than social scores (*M* = -0.20, *SE* = 0.07) for the low SES group [*Mean diff* = -0.32, *p* = .01, CIs 95%[-0.57, -0.07]; Figure 2.12. Social latent scores were significantly higher for the high compared to low SES group; *Mean diff* = 0.25, *p* = .002, CIs 95% [0.09, 0.40]; non-social scores did not differ significantly between the two SES groups; *Mean diff* = -0.12, *p* = .199, CIs 95%[-0.30, 0.06].

Table 2.10: Mean, standard deviation and N for social and non-social latent scores for high and low cluster groups and the whole sample

Latent score condition	Cohort	High SES			Low SES			Whole sample		
		M	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*
Social	Term	0.03	0.67	447	-0.13	0.66	119	-0.001	0.67	556
	Preterm	0.05	0.70	117	-0.28	0.66	29	-0.02	0.70	146
	Total	0.04	0.67	564	-0.16	0.66	148	-0.004	0.68	712
Non-social	Term	-0.05	0.80	447	0.13	0.76	119	-0.01	0.80	566
	Preterm	0.05	0.80	117	0.10	0.70	29	0.06	0.78	146
	Total	-0.03	0.80	564	0.13	0.75	148	0.002	0.79	712

Figure 2.12: Mean scores for social and non-social latent scores for high and low SES cluster groups



2.3.3.3 Differences in individual eye-tracking measures across SES groups

To investigate whether performance across eye-tracking tasks differed across SES clusters, multivariate ANOVA (MANOVA) was performed with SES cluster group as an independent variable with two levels (high and low SES). Table 2.11 lists the eye-tracking metrics which were included as dependent variables; these were the same as those used in the structural equation model previously. Box's test was not significant at the p < .001 level [*Box's M* = 405.18, *p* = .01]. A significant main effect of SES was found [*F*(24, 395) = 1.80, *Hotelling's Trace* = 0.11, *p* = .01, η_p^2 = 0.10]. Significant differences between SES groups were found for the mean proportion of looking time to faces in the pop-out task [*F*(1, 418) = 5.06, *p* = .03, η_p^2 = 0.01], proportion of trials that were correctly anticipated [*F*(1, 418) = 5.30, *p* = .02, η_p^2 = 0.01] and reaction time [*F*(1, 418) = 6.57, *p* = .01, η_p^2 = 0.02] during the pre-switch phase in the reversal learning task, and mean reaction time for trials in which participants were correct in the working memory task *F*(1, 418) = 5.94, *p* = .02, η_p^2 = 0.01]. Means indicated that the low SES group looked more to faces in the pop-out task, were less likely to correctly anticipate trials and were slower in the reversal learning task and were slower in the working memory task (see Table 2.11, Figure 2.13).

Figure 2.13: Mean standardised scores for key eye-tracking variables for high and low SES groups (stars denote a significant difference)

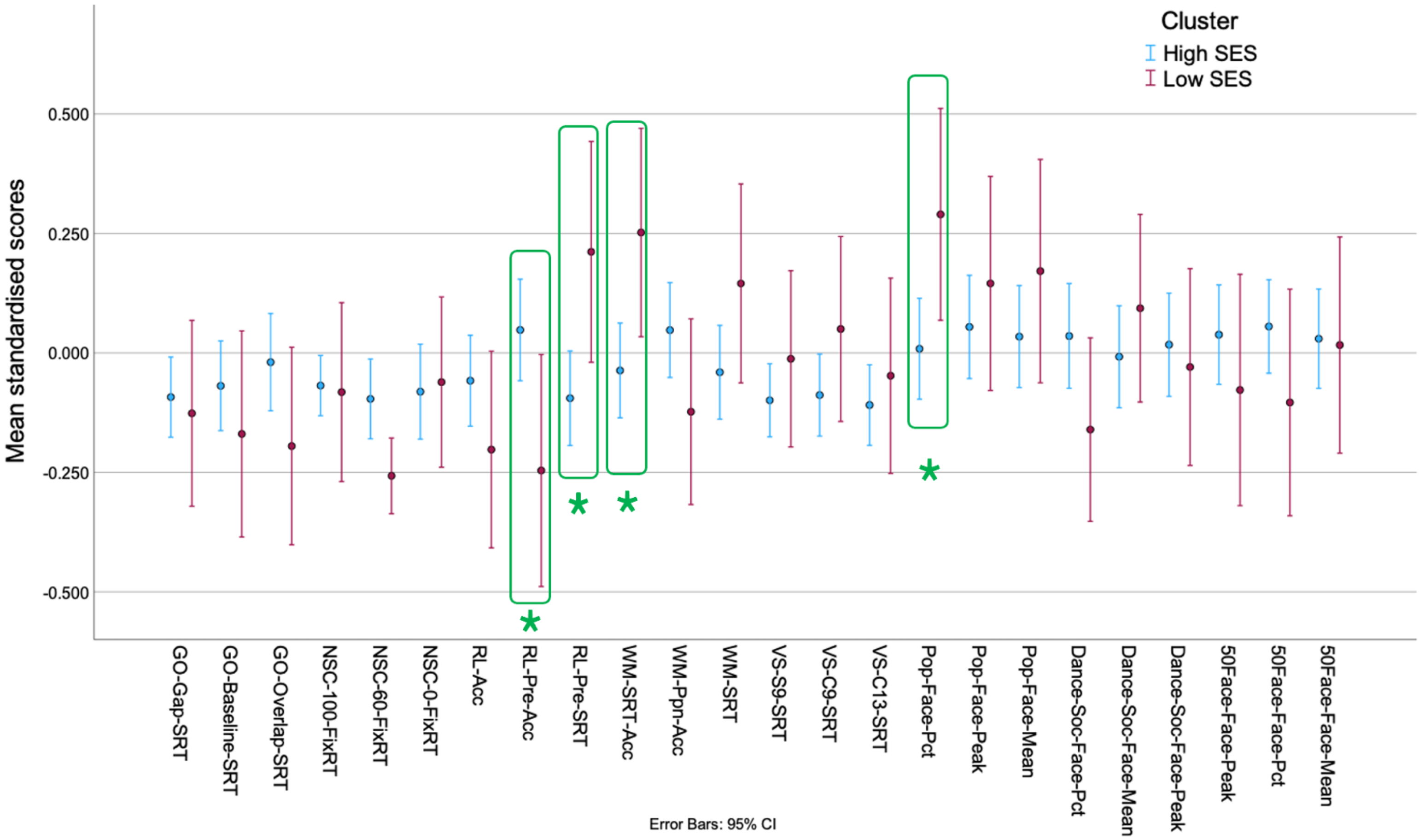


Table 2.11: Mean, standard deviation and N for key eye-tracking variables for high and low SES groups

Task	Variable	High SES (*N* = 343)		Low SES (*N* = 77)	
		M	*SD*	*M*	*SD*
Gap-overlap	GO-Gap-SRT	5.53	0.10	5.53	0.11
	GO-Baseline-SRT	5.73	0.12	5.72	0.13
	GO-Overlap-SRT	5.89	0.18	5.86	0.17
Non-social contingency	NSC-100-FixRT	660.22	370.17	651.78	514.46
	NSC-60-FixRT	494.08	221.63	448.69	98.20
	NSC-0-FixRT	519.75	232.37	524.74	195.48
Reversal Learning	RL-Acc	0.63	0.15	0.61	0.15
	RL-Pre-Acc	0.71	0.27	0.64	0.29
	RL-Pre-SRT	0.67	0.19	0.73	0.20
Working Memory	WM-SRT-Acc	707.45	207.12	771.46	212.87
	WM-Ppn-Acc	0.51	0.18	0.48	0.17
	WM-SRT	688.84	168.78	722.81	167.62
Visual Search	VS-S9-SRT	958.49	249.19	988.53	281.07
	VS-C9-SRT	1361.06	440.49	1436.39	464.48
	VS-C13-SRT	1503.19	449.23	1537.88	460.47
Face pop-out	Pop-Face-Pct	0.31	0.13	0.35	0.13
	Pop-Face-Peak	1.16	0.49	1.21	0.47
	Pop-Face-Mean	0.78	0.28	0.81	0.29
Dancing Ladies	Dance-Soc-Face-Pct	0.21	0.14	0.18	0.12
	Dance-Soc-Face-Peak	0.64	0.38	0.63	0.34
	Dance-Soc-Face-Mean	0.18	0.08	0.19	0.07
Fifty Faces	50Face-Face-Pct	2.27	1.09	2.14	1.19
	50Face-Face-Peak	0.55	0.12	0.53	0.14
	50Face-Face-Mean	0.41	0.17	0.41	0.17

2.4 DISCUSSION

Although much existing research has used eye-tracking to assess visual attention in infants and young children, little work has considered the feasibility of conducting large-scale eye-tracking studies in toddlers. In this study, data were collected from a battery of eight eye-tracking tasks completed by over seven hundred 18-month-olds. Data was successfully collected from 97.9% (697) of toddlers on at least one task. Analyses found expected condition effects in seven out of eight eye-tracking tasks, with only the working memory task not finding the expected effects at the group level. A hypothesis-driven SEM provided a good fit to the data, indicating that in addition to indices from individual tasks, the battery can be used to extract global measures of social and non-social attention. Using maximum likelihood imputation allows latent variable measures to be extracted for all participants in the sample. A cluster analysis across several SES variables found two clusters which might be thought to represent a low and a high SES group. Analyses found group differences in performance on certain eye-tracking measure across these two clusters: specifically in relation to looking to faces and cognitive control. Overall, the current eye-tracking battery provides a feasible and objective measure of visual attention in 18-month-old children and provides insights about how profiles of visual attention may differ in relation to previous experience.

2.4.1 Overall feasibility

The findings indicated that overall feasibility of the study was good. For every task at least 70% of children provided enough data to be included in the analyses; this was 90% for four of the first five tasks in the battery. If cut-off criteria were applied, at least 64% of the whole sample would be retained; over 86% for the four tasks with highest retention. Fewer than 10% of children failed to meet cut-off criteria, indicating that most participants who completed the task produced valid data and further supporting that use of the eye-tracking battery with 18-month-olds is feasible. Furthermore, conditions effects were the same for all but one condition comparison before and after children who did not meet typical cut-offs were excluded, indicating that patterns of performance reported in this age range do not just capture effects seen in children with stronger attentiveness.

Despite validation procedures, data quality still impacted extracted metrics. Metrics which were associated with either accuracy or precision metrics at a level of >0.2 (suggesting they should be controlled in research using these tasks) included gap reaction time in the

Gap-Overlap task, reaction times in the non-social contingency task, singleton nine visual search accuracy, percentage looking to faces and multiple peak look duration variables in the dancing ladies and percentage looking and peak look duration to faces in the Fifty Faces free viewing video. This is likely because even with expanded AOIs, lower accuracy and precision leads to less certainty in gaze sample classification. Loadings in the model suggested precision and accuracy measures may be better captured as one data quality factor, with median lost samples another; as such, controlling for these factors may be the most parsimonious approach.

2.4.2 Task robustness

Most tasks in the current battery produced robust condition effects in the expected direction. These included competition effects in visual orienting, the ability to use memory to find a video, set size and pop-out effects on search for a visual target and the preference for faces in early development. These findings indicate that it is possible to robustly measure many previously observed effects, even within a large battery of tasks. This has important implications for the replicability and validity of these findings, strengthening support for their generalisability and robustness. In the light of increasing concerns about the reproducibility of cognitive psychology findings (Huber et al., 2019), this set of replications provides important reassurance that developmental science has produced a range of robust observations about the developing infant attention system.

The one exception was the working memory task, which provided no (or very weak) evidence of working memory at the group level when trial-level reaction time data were analysed. This may reflect the fact that this skill is relatively fragile at this age; indeed, there are few robust demonstrations of single-trial level working memory success at 18 months (Hendry et al., 2016). One important task feature may be that the child had to actively ‘find’ the object with their gaze, rather than passively watch sequences of objects as has previously been used in other working memory tasks with younger infants (Ahmed & Ruffman, 1998; Baillargeon et al., 1985). Indeed, Hood et al. (2003) found that the same 2.5-year-olds demonstrated ability to recognise impossible locations of an object but failed to complete active search and retrieval of the object, indicating a disparity between active search and passive observation. Although children typical succeed at passive visual tasks at a substantially younger age than behavioural analogues (Ahmed & Ruffman, 1998; Baillargeon & Graber, 1988; Hofstadter & Reznick, 1996), it may be that this active component makes tasks more difficult. This additional difficulty when actively searching for the object may have caused some toddlers to fail in the current study, leading to the observed null findings for this task.

2.4.3 Structure of visual attention

A hypothesis-driven SEM provided a good fit to the data once strong within-task associations were accounted for. The most parsimonious structure indicated two latent variables that can be interpreted as social attention (interest in people) and non-social attention (speed and accuracy of saccades). Though the current tasks did enable condition difference scores to be calculated, these can be unstable (combining the noise inherent in both variables), therefore measures of accuracy and reaction time were used in this model. Indeed, Draheim et al. (2019) proposed that the lack of correlation between attention measures in many studies is due to methodological issues with attention capture tasks, rather than it being that attention is not a unified concept. They conclude that accuracy-based measures may provide more reliable and valid measures of attention control than reaction times and difference scores. This assessment was based on several judgements, including how well metrics from various tasks inter-correlated with one another; notably, considerable correlations were found even between metrics from tasks which make markedly different demands on the participant. Reaction times were used in the model to help with comparison to previous literature and chose accuracy over condition differences scores to ensure greater validity and reliability of this work.

The lack of correlations between the two latent variables in our SEM is consistent with models in which social attention represents a distinct construct. The large sample size in the current work implies this result is not due to lack of statistical power and supports the notion that metrics do not associate. Based on this, eye-tracking metrics should be carefully chosen in future work in this field, to ensure that tasks measure an appropriate facet of visual attention. Although many models of attention distinguish exogenous and endogenous orienting, inclusion of this distinction in our modelling did not explain significantly more variance than a two-factor model. Neuroimaging studies have indicated that both endogenous and exogenous attention shifts are mediated by the same large-scale fronto-parietal networks, indicating they may be closely related constructs (Peelen et al., 2004). However, other studies have shown distinct impacts of environmental variation (such as screen exposure) on exogenous and endogenous attention shifting

(Portugal et al., 2021). A model incorporating the endogenous/exogenous distinction did provide an adequate fit to the data and may be preferred by investigators who wish to distinguish these components of visual attention at the individual level.

2.4.4 Relation between socioeconomic status and visual attention

A data-driven approach to grouping SES found two clusters which mapped to a generally high and a low SES group. This approach enabled SES to be operationalised as a single variable without reducing individual measures into a single composite score. That is, analyses considered whether (and how) profiles differed across multiple SES measures, therefore allowed different SES measures to have different scores within a grouping. For example, clusters could have revealed two different SES profiles which did not map onto typical high and low SES groups, but instead one which had higher education and CSPS score and another with lower education and lower IMD scores (i.e. indicating lower deprivation). Both such groups would have a mix of typically high and low SES indicators and individual SES measures would need be considered to characterise such groups. That the clusters found here do map to typical SES characterisations fits with previous work suggesting associations between different SES measures, though the varying distances between measures (i.e. significant differences in education scores but large difference in IMD scores), supports that measures are not completely interchangeable (Braveman et al., 2005). Measures may capture different structural features of an individual’s environment, with experience of particular SES features being more (or less) likely to lead to experience of other conditions (Antonoplis, 2023).

Considering whether profiles of visual attention vary according to key components of visual attention revealed different patterns of performance dependent on SES grouping. Latent scores relating to social and non-social attention did not differ for the high SES group, however the low SES group showed relatively lower social and higher non-social scores. Further investigation revealed that the higher SES group looked less to faces, were faster to find a hidden object in a working memory paradigm and were faster and more accurate in the learning phase of a cognitive control task compared to the low SES group.

Longer looking to faces has previously been associated with atypical development (Elsabbagh et al., 2013; Hendry et al., 2018; Rose et al., 2001), with some researchers suggesting this behaviour indicates an overly focal attention style (Elsabbagh et al., 2013). Interestingly, an SES-related difference in face looking was only found for the percentage of looking to faces, but not for peak or mean looking duration, in the current study. Though the reasons for this are a little unclear, it could suggest that children in the low SES group tended to look back to the face more frequently than the high SES group, though individual looks were not longer. In this way, the pattern observed here may differ from a disengagement difficulty and instead indicate a visual foraging technique which favours faces over other stimuli. In fact, a pattern of more attention to faces has previously been associated with better positive affect (Sheese et al., 2009), emotion regulation (Morales et al., 2005) and theory of mind (Wellman et al., 2008), whilst decreased attention to people has been interpreted as a delay in social development (Clearfield & Jedd, 2013). Given such, the more looking to faces by children in the low SES group found here should perhaps be considered as a strength associated with lower SES and could have implications for an adaptive view of how experiences impact SES.

In addition to specificity in the measure, more looking to faces was only found in the pop-out task which uses static images but not in free viewing tasks which utilise dynamic videos. This is line with other studies which have found different results in different types of task (Chevallier et al., 2015; Del Bianco et al., 2022) and may be due to different processing requirements for static versus dynamic stimuli. Videos provide a lot of exogenous motion cues for attention which static images cannot, therefore it is possible the current findings were driven by differences specifically when such information is limited. There were additionally differences in trial design, with the pop-out task using a trial-based design with short trial durations, whilst other relevant tasks utilise a free viewing design with longer durations, which could have impacted results. Taken together, these factors and the current findings could suggest a relation between SES and the ability to process simple social information when there are shorter viewing times only (i.e. not when there may be broader input). This might indicate the existence of different patterns of reliance on stimulus features in environment-driven attention which may be related to children’s early experiences.

The findings here that children in the low SES group had reduced working memory and cognitive control abilities fits with a wider body of evidence suggesting that SES impacts these domains. In particular, low-SES infants showed reduced cognitive control compared to high SES infants in an A-not-B reaching task (Clearfield & Niman, 2012). This used a similar design to the reversal learning task here, whereby there was a learning phase followed by a post-switch phase in which the animated side was changed, thus supporting that environmental experiences typically associated with low SES may impact young children's ability to control their attention. The current findings relate to the pre-switch phase of the task, in which children were required to inhibit looks to the side of the screen they chose initially and learn to anticipate the appearance of a visual reward, with the low SES group showing both reduced speed and accuracy. Poorer inhibitory control has been suggested by other authors to be the result of a cognitive trade-off in which shifting between tasks is enhanced at the cost of inhibitory ability, and which may be particularly useful for succeeding in unpredictable environments with fewer opportunities (which are generally more typically in low SES contexts) (Ellis et al., 2017; Mittal et al., 2015).

A similar pattern has been suggested in relation to working memory, with evidence that this ability may be reduced, whilst procedural learning is enhanced, through experience of poverty (Dang et al., 2016). Other work has found mixed results regarding the relation between SES and working memory, with Ellis et al. (2017) summarising that some aspects of working memory may be enhanced whilst others are reduced in response to experiences. The current study found that the low SES group was slower to correctly find the hidden object though no differences in accuracy were found. This may indicate that processing speed is specifically impacted by experience, which is similar to findings by McCoy et al (2015), though their results were in the opposite direction and related to different cognitive domains.

As we have seen in relation to working memory and cognitive control, an adaptive-based approach to the relation between SES and visual attention may help to explain the results found here, though more work is needed to investigate this. Existing literature in this area has used a variety of tasks, SES measures and participant samples, meaning it is possible that findings relating to trade-offs between different attentional abilities may be explained by measurement disparities or developmental changes. Nonetheless, findings within the current study may be considered together to show increased focus to static faces, reduced working memory speed and reduced cognitive control speed and accuracy.

2.4.5 Limitations

Though this work draws strengths from its large sample size and substantial task battery, there are some limitations. Whilst it is assumed that eye-tracking tasks provide a direct measure of attention, assumptions are made about specifically what aspect of attention tasks are tapping. Most tasks in the current battery have been widely used previously; the current study serves to establish previously found condition effects in this large toddler sample. Yet, inaccurate assumptions about the processes captured by different tasks may lead to mistakes in how effects are interpreted. Additionally, processing of data has been carried out in line with other work using these tasks, but this could be problematic. Measures calculated directly from the eye-tracker remove an element of subjective assessment, but there remain numerous options for processing (i.e. selecting size of area of interest (AOI), defining data validity, etc.); some of these decisions may impact the results found in later analyses. Such limitations ought to be considered by all eye-tracking researchers. Indeed, there is a continuing effort to establish reliable and valid eye-tracking batteries in young children (van Baar et al., 2020); the current work contributes to this field by establishing the robustness of these eye-tracking tasks in 18-month-olds. That said, the current study only tested a very specific age and further work spanning development will help establish robustness of this battery over toddlerhood.

The SES measures used here could also be a confine of this work. Whilst the statistical approach used here did not reduce individual SES measures into a single composite score, decisions were nonetheless made about which measures to include in analyses. Education and occupation were chosen in line with much previous work, whilst IMD scale and CSPS score were included as additional measures of a child's environment. The IMD scale is determined from a participant's postcode, meaning it could provide a measure of deprivation which differs from individual's own specific SES level. Additionally, whilst SES measures used here were gathered at different time points, changes over time are not considered in the current analyses, meaning potentially time-sensitive impacts on development may be overlooked. Gathering both SES and visual attention information from children longitudinally may be additionally informative about the development relations between the two and could contribute to understanding about causal relations between early experiences and cognition.

2.4.6 Conclusion

The current study demonstrated that data-derived groupings can be used to investigate SES-related differences in performance on attentional measures, with evidence that profiles of visual attention differed in relation to SES experiences in early life. Specifically, latent scores relating to social and non-social attention did not differ for the high SES group, however the low SES group showed relatively lower social and higher non-social scores. Further investigation revealed that the higher SES group looked less to faces, were faster to find a hidden object in a working memory paradigm and were faster and more accurate in the learning phase of a cognitive control task. Such information could be of great use in understanding what support may help individuals with poorer attention skills. This is particularly true for children from poorer, low-resource backgrounds, for whom understanding attentional adaptations may be beneficial both for developmental and societal support. The current results could also be used as a basis for neurocognitive work linking together eye-tracking and imaging information, which may be facilitated by its role as part of the wider DHCP project (http://www.developingconnectome.org/).

In addition to novel findings about the relation between early experiences and profiles of visual attention, this study showed that a large eye-tracking battery can be successfully used with 18-month-olds and can produce both task-based and cross-task indices that may be useful for large-scale assessment of visual attention in toddlers. It provides a strong foundation for future work which may investigate differences between different groups of toddlers, including typically developing children and those with clinical conditions. The large amount of data included in this study lends itself to studies considering individual differences in attentional processing. This includes measuring individual differences in the underlying structure of visual attention as well as with task effects. As demonstrated in this study, understanding how the structure of attention differs across infants could provide insights which a task-based approach cannot, including possible environmental impacts on the underlying structure of visual attention.

There is large potential for the present work to contribute to pioneering research; it is a blueprint for eye-tracking research with toddlers which may lead to improved theoretical understandings about cognitive functioning in toddlerhood.

3 TOOLS IN THE REAL WORLD: ALPHA AND THETA RELIABILITY IN OLDER TODDLERS

Abstract

Optimisation of brain health could have an important impact on individual's health and well-being (World Health Organization, 2022), particularly for those where development may be sub-optimal due to early environmental experiences. Neural methods in particular hold great potential for improving current understanding of how early experiences impact development and may provide mechanistic insights about this relationship which cognitive and behavioural methods cannot. Much work has made use of neuroimaging methods (such as EEG) for measuring neurodevelopment in infants and preschool children, though these have typically taken place in highly controlled labs and with a highly skewed sample of the population. Children from families of lower socioeconomic status have tended to be missed out of this research, meaning findings may not be robust and generalisable to the general population. Development of portable and wireless systems now provide tempting potential to begin to overcome some of these difficulties, by providing the possibility of taking research to the field. This may facilitate a more diverse sample of toddlers to be included, enabling research to be more representative and generalisable. Before such methods can be used to make novel findings about neural development, it is important that measures are first assessed for psychometric properties such as reliability.

This chapter utilised a two-visit design to assess the reliability of a wearable, wireless Enobio EEG system with typically developing toddlers. It focussed specifically on alpha and theta power measures during video viewing, as these are thought to be involved with cognition and learning and may be particularly involved in aspects of development which are impacted by early environmental experiences. Numerous measures of theta and alpha power were calculated, with significant differences in EEG power found between different frequency bands, brain regions, video conditions and other comparisons. Reliability for these measures varied considerably, with relative theta power over the whole head and whole video providing the most reliable measures; researchers are recommended to carefully consider reliability values when choosing measures in future studies.

Situate in thesis

This chapter fits into the wider aim of this thesis to improve data collection methods across a diverse sample of toddlers, which may then be scaled up for use in larger studies. This study utilised a portable neuroimaging system, which holds potential for including a broader range of children in developmental cognitive neuroscience research as it is well-suited to use in field or community settings, thereby reducing some travel burdens. Additionally, the specific neural measures considered in this chapter appear related to aspects of early cognitive development which may be particularly impacted by early experiences, meaning they may be especially useful focusses for optimisation of development. When working towards improved methods for measuring cognition and developing support, it is important that measures are reliable. The current chapter addressed this, using a test-retest study to assess the reliability of a neurocognitive measure using the 'Braintools' wireless toolkit.

3.1 INTRODUCTION

Brain health refers to the level of neural functioning across various cognitive, behavioural, and other domains (World Health Organization, 2022). Optimising brain health could have significant impacts both for individuals' health and well-being, and more broadly for society (World Health Organization, 2022). This may be particularly true for individuals who have not reached their developmental potential due to experience of adversity in their early years of life. To optimise brain functioning, a thorough understanding of optimal and sub-optimal brain development, using large-scale, adaptable tools, is necessary. Whilst advancements have been made, there remains much to be done to develop effective methods for measuring brain development during the early years of life. This includes expanding the age range of participants typically included in neuroimaging research, as well as reducing the cost and improving the scalability of research. It is important to adapt neuroimaging methods to meet these needs, since a better understanding of brain mechanisms may reveal crucial insights into children's development that cognitive and behavioural measures do not. For example, neural differences may be apparent before behavioural differences are, meaning early markers of later abilities could be found which may enable children at risk of poor outcomes to be identified before difficulties are so profound. This may allow interventions to be more effective, both because they may occur earlier and because they may be more specific to the learning mechanism underlying behavioural development. There is thus a need for the development of scalable, portable, and low-cost neuroimaging kits and research set-ups which can overcome some of the difficulties of current research.

3.1.1 Portable neuroimaging: electroencephalography

One neuroimaging method which holds such potential is electroencephalography (EEG). EEG measures oscillations in neural processing involved with information consolidation in learning and memory. It is non-invasive and dynamic, enabling the brain activity of young children to be recorded during natural behaviour and permitting real-time access to the neural rhythms underpinning information consolidation. EEG oscillations can be segmented and separated into frequencies and regions, leaving a specific signal which may be linked to a particular cognitive function. Frequencies are typically organised into frequency bands, with average EEG power calculated over a particular range of frequencies: increasing from lower to higher frequencies, frequency bands commonly used are delta, theta, alpha and high-frequency activity of beta or gamma.

Recent technological advancements have led to the development of portable and wireless EEG systems. Not only do such systems enable young participants more freedom of movement, but they also have advantages for the potential use of EEG in more low resource and/ or field settings. Current research typically happens in labs with relatively educated, high-income families. As has been discussed extensively in chapter 1 of this thesis, a large majority of children at risk of poorer-cognitive outcomes live in lower resource settings, yet current neurocognitive research tends to include samples of participants which are biased towards higher resource settings, likely driven by the resources necessary for families to participate in this research. There is a pressing need to develop new ways to expand this research into the community. Portable EEG systems could help take neuroimaging research to the field, including settings in other countries, particularly low-income countries where resources are particularly limited, or in lower socioeconomic status (SES) communities within the UK (see Troller-Renfree et al., 2021 who have achieved similar in the US). Increasing representation within neuroimaging research would enable research to be more generalisable and may be important for future identification of sub-optimal brain development.

As outlined by Bhavnani et al. (2021), in addition to conducting research outside of highly controlled lab settings, there is also a need to find robust measures for low-density (i.e.

fewer electrode) neural set-ups, to avoid the use of small, homogenous samples in studies, and to standardise methods and best research practice (p24). Low-density set-ups may be particularly beneficial for research with toddlers, since using fewer electrodes reduces capping time and may increase tolerance of the EEG cap. Systems which are low cost and with a simple set-up may further increase accessibility of neuroimaging research for smaller, less wealthy/ well-established settings, thereby facilitating large scale neuroimaging studies in low resource settings. One study has already demonstrated how a low-density (20-channel) mobile EEG system was used to collect neural data from a large sample (N > 400) of infants from low-income backgrounds in the United States (Troller-Renfree et al., 2021). Researchers used their experience during this project to develop methodological and analytical guidelines for how high quality EEG data can be collected from participants' homes, which can help other researchers make use of mobile EEG systems for similar purposes.

Whilst there is an obvious need for the development of scalable, portable, and low-cost neuroimaging systems and research set-ups which can overcome some of the difficulties of current research, it is important that psychometric properties of measures are first assessed before being used to make novel findings about neural development. Psychometric properties are essential to ensure measures - and therefore findings that relate to them - are accurate, meaningful, and consistent. Psychometric properties include validity and reliability, with a measure considered valid if it measures what it is supposed to measure and reliability referring to the consistency of a method; whether it can repeatedly gather the same measures under the same conditions. Whilst assessing reliability of measures is important when considering group differences, it is particularly critical for work which considers individual differences (Cooper et al., 2017). The impact of measurement error can be minimised by increased samples in group designs, which cannot be done when using measures at the individual level (Cooper et al., 2017). Reliability, as a measure of the degree to which measurements are free from error, is therefore particularly important for individual differences research. Of note, the reliability and validity of an EEG measure might be impacted by factors including participant age and experimental setting; that is, the psychometric properties of a given measure may differ across contextual uses. It is therefore important that these are assessed for the particular population and context of interest (Cooper et al., 2017).

The current study utilised a two-visit design to assess test-retest reliability and feasibility of a low-cost, low-density, portable EEG system for use in field settings. The project used a toolbox – *Braintools* (also reported here: (Haartsen et al., 2021)) – which included a range of visual and auditory tasks which are commonly used in developmental neuroscience research. This chapter focussed on differences in EEG power during a visual task. Much work has investigated how different EEG rhythms may be associated with cognition; in the current chapter measures of power in both the theta and alpha range were considered.

3.1.1.1.1 Frequency bands

Before reviewing this work in more detail, a note should be made about frequency ranges and how these are banded together and labelled, as these differ across ages and studies. Though frequency bands are well-delineated in adults, these are typically not similarly applied to young children, given the evidence suggesting that EEG in the first years of life occurs at a much lower frequency than in adults (Bell, 1998; Lindsley, 1939; Saby & Marshall, 2012). Studies have investigated theta and alpha from infancy to preschool, yet it is unclear how frequency bands should be defined across these ages. Stroganova and Orekhova (2007) outlined that in infants, theta is generally considered across the 4-6Hz band whilst alpha is between 6-10Hz, though (Saby & Marshall, 2012) note that 3-6Hz is the most commonly used infant theta band and 6-9Hz is generally considered comparable to adult alpha. Whilst ranges are at least somewhat defined for infants, there is less clarity for toddlers and preschoolers and frequency bands vary across studies. In a rare study with toddlers (2-year-olds), Cuevas et al. (2012) considered theta as 3-5Hz and alpha as 6-9Hz; whilst in preschoolers theta has been calculated over ranges including 4-8Hz (mean age 5 years, 5 months) (Orekhova et al., 2006) and 3-6Hz (mean age 52 months) (Meyer et al., 2019), and alpha over 6-10Hz (mean age 4.59 years) (Begnoche et al., 2016). Whilst 6-9Hz has been largely considered an appropriate band for alpha power in infant work, (Marshall et al., 2002) found that many of the 51-month-old participants in their study showed spectral peaks at 9Hz, thus they recommended an extended alpha band of 6-10Hz for this age.

Given that frequency ranges are predominantly defined by how they function, it is plausible that subtle differences in what appears as theta and alpha may differ across different populations and with different experimental designs, as well as over developmental age. In line with much developmental work and aligned particularly with Cuevas et al. (2012), the current work calculates theta over 3-6Hz and alpha over 6-10Hz, with this slightly higher upper limit of alpha ensuring alpha can be captured for the oldest children in this study (as per Marshall et al., 2002).

3.1.2 Theta and alpha power as neural measures of cognitive development

Theta and alpha power are both thought to be involved in various cognitive processes, with work linking alpha oscillations to visual attention and inhibition (Foxe & Snyder, 2011; Klimesch et al., 2007; Klimesch, 2012), and theta activity with social attention, processing of emotional information and memory (Guderian et al., 2009; Z. Jiang et al., 2017; Jones et al., 2015). Given their relation to functions such as attention and executive functioning, which are important processes underpinning many other cognitive skills and which show significant development from infancy into childhood, it follows that alpha and theta oscillations may, too, be developmentally significant (Cellier et al., 2021). Moreover, measures of theta power in the first years of life have shown promise for predicting later cognitive ability (Braithwaite et al., 2019; Jones et al., 2020), thus may hold particular potential for the development of support to optimise brain function.

3.1.2.1 Alpha

Early developmental EEG work focussed on alpha rhythms (Berger, 1929), with initial investigations focussed on establishing what constitutes alpha in young children. Two main findings relating to the 6-9Hz frequency band in infants and toddlers indicate that this band is analogous to adult alpha, which is typically considered between 8-12Hz . First is that, in adults, the dominant EEG rhythm is generally considered to be in the alpha frequency range (Klimesch, 1999), whilst in longitudinal developmental work, 6-9Hz has been found to be the most dominant frequency from 5 months to 4-years-old (Marshall et al., 2002). Secondly, adult alpha is characterised by amplitude increases in occipital regions during a baseline period; the same pattern was found over the 5.2-9.6Hz frequency range in infants during what’s considered to be an equivalent baseline period (Stroganova et al., 1999).

Having established this frequency range as alpha, work has since investigated different alpha measures, and their relation to cognition. In doing so, several key alpha rhythms have been discovered. In infants, a rhythm in the alpha range has been found over central regions, which seems to emerge at around 4- to 5-months-old (Marshall et al., 2002; Smith, 1941). This is believed similar to adult mu (Smith, 1939; Stroganova et al., 1999) and thought to be linked to action processing (Marshall et al., 2011; Southgate et al., 2009b). A different, posterior alpha rhythm has also been found in young children (Marshall et al., 2002; Stroganova et al., 1999). This posterior alpha appears to be associated with information processing, and may be particularly linked to inhibition (Klimesch, 2012).

Early alpha work found that visual input modulated posterior alpha in adults, with higher alpha amplitude when eyes were closed and lower amplitude when eyes were open (Berger, 1929). Similar patterns have been found in infants, with maximal amplitude found in occipital regions during a condition of darkness, which was significantly larger than when children were visually attending to bubbles (Stroganova et al., 1999). This reduction in amplitude/ power (power being amplitude squared) in response to visual stimuli appears in contrast to the assumption that the magnitude of neural response is reflective of the level of processing required and it has thus been suggested that this pattern reflects engagement of inhibitory processes (Klimesch, 2012). Specifically, it is thought that higher amplitude/ power might reflect active inhibition of certain neural networks which are irrelevant to the specific processing required for a given task, such that attention to the task can be maintained. For example, work has found that alpha power was larger over visual regions when the task required attention to auditory input (Foxe et al., 1998) and was larger over parietal regions (usually implicate sensory processing) when the task required the ventral stream (i.e. involved in visual perception) (Jokisch & Jensen, 2007). In infants aged 8-11 months, greater amplitude in posterior alpha was found to be related to longer bouts of attention during the anticipatory period of a game of peek-a-boo (Orekhova et al., 2001). Given the role of posterior parietal networks in processes to shift attention (Posner et al., 1987), the authors suggested that this finding might reflect active inhibition of certain parietal networks which is required to prevent attention shifting from the centre to the peripheral visual field (Orekhova et al., 2001). Similar findings have been reported in other domains, with evidence of higher alpha power in parietal regions during the anticipation of movement (Westphal et al., 1993) and when short term memory demands are high (Klimesch et al., 1999). Despite differences between cognitive domain of interest and task design, parietal alpha was specifically found in relation to focussed periods of attention to the centre of the visual field in these studies, thereby might further support that alpha is implicated in maintaining this focussed attention (Orekhova et al., 2001). In this way, alpha power may be important for inhibition of task-irrelevant information and facilitation of attention to relevant stimuli. In fact, Klimesch (2012) argued that the cognitive role of alpha is not in relation to any specific domain but to broader functions which underpin many other cognitive processes.

Although alpha is generally the most-studied frequency band range, there is still somewhat limited research involving toddlers or young children. In a study involving 2-year-olds, three electrode pairs in frontal, parietal and occipital regions showed higher alpha power (6-9Hz range) in a condition involving memory encoding compared to baseline (passively watching a video) . For frontal and parietal regions, power was also higher during memory retrieval compared to encoding. There were additional group differences in alpha power in relation to behavioural recall performance, with higher power relating to better performance. This might suggest that alpha power is related to cognitive development, which is in line with work suggesting that alpha power increases over the course of brain maturation.

In fact, it seems that relative alpha power increases and theta power decreases over developmental age (Clarke et al., 2001; Marshall et al., 2002). Though in adults the dominant EEG rhythm is generally considered to be in the alpha frequency range (i.e. Klimesch, 1999), in young children it appears to be within the theta band, which some have suggested may reflect a brain that's prepared for optimal synaptic plasticity (Stroganova & Orekhova, 2007). Over all scalp regions, the dominant frequency appears to increase with age (Cellier et al., 2021; Marshall et al., 2002), though different studies report different ages at which the dominant frequency shifts from theta to alpha. For instance, Cellier et al. (2021) suggest that this shifts occurs at around 7 years, whereas Marshall and colleague's (2002) work indicates a shift may be apparent at around 10-months-old. The current study may shed light on some of this uncertainty, by comparing relative power in theta and alpha in a sample of participants between these two ages. In addition, it seems that alpha oscillations may be linked to attentional control, therefore this chapter will look at whether alpha power varies across video viewing and depending on stimulus type.

3.1.2.2 Theta

Research investigating theta power has used various measures, including average power over a time period and, more recently, metrics of power change over time. Theta oscillations have been linked extensively to memory, with average theta power having been linked to memory encoding (Guderian et al., 2009; Sederberg et al., 2003), memory accuracy (Crespo-García et al., 2016) and visual working memory (Pavlov & Kotchoubey, 2022). In particular, one infant study found that frontal theta power during object exploration was significantly associated with subsequent recognition of those objects (Begus et al., 2015). Indeed, it is thought that frontal and temporal theta rhythms may be driven by hippocampal theta, thus further emphasising this association (see Lega et al., 2012). Given that theta relates to memory encoding and formation, it can also be said that theta is implicated in learning (Begus & Bonawitz, 2020).

Outside of the memory domain, higher average theta power has been found in infants during novel object exploration (Orekhova et al., 2006), during anticipation of peek-a-boo (Orekhova et al., 1999) and in conditions involving social stimuli (Orekhova et al., 2006); all conditions which are likely to involve learning for participants of this age. In one study, authors found higher average theta power during a condition involving incorrect puzzle configurations (i.e. error detection) compared to a correct configuration condition in both toddlers (mean age 16.75 months) and adults, with the degree of difference in theta power between conditions additionally associated with SES as measured by a composite of parental education, parental occupation and family income (Conejero et al., 2018). Average theta power has also been associated with duration of look to a toy, with the same study also finding that continuous oscillations in theta power predicted whether infants visual attended to a toy, particularly during solo play (Wass et al., 2018). These authors further found a similar association between EEG power and visual attention in adults, though this was found at a higher EEG frequency in the alpha band (6-12Hz). This finding of analogous functions of theta power in infants and alpha power in adults is in line with other research indicating a potential shift from dominance of activity in the theta band shifting towards the alpha band over development (Marshall et al., 2002).

In addition to comparing average theta power between conditions, some work has considered measures of how much theta power changes over time. In a study involving preschoolers (mean age 52 months) (Meyer et al., 2019) found that frontal theta power increased over multiple presentations of a fixation cross during rest periods of a cognitive task. Though no task information was presented during these periods, in the latter two rest periods participants were required to maintain attention, encode information, and prepare a response to the task whilst in the first period the task had not yet begun. Thus, theta power appears to be increase with greater task engagement and increased cognitive control. This is in line with other work which has found an increase in theta power during anticipation of information processing. For example, increases in theta activity prior to presentation of stimuli have been associated with subsequent recall rate (i.e. indicating encoding) (Guderian et al., 2009), with some work also finding a similar pattern for alpha activity (Fell et al., 2011). This phenomenon has also been observed in social contexts, with theta increases found during anticipation of information from an adult, but before information was given (Begus et al., 2016). This was further modulated by the informative potential of the adult, with theta power higher when infants were expecting information from a communicative compared to an uninformative adult. Such findings suggest that increases in theta power may indicate a preparatory learning state (Begus et al., 2016) and could be reflective of cognitive engagement.

Other findings also support that theta power may be related to processing of social information, with higher theta power when observing social versus non-social stimuli (Jones et al., 2015). Despite this finding that average theta power is higher when viewing social compared to non-social stimuli, work which also looked at the *change* in theta power over the course of during social and non-social stimuli viewing found no condition difference (Jones et al., 2020). This demonstrates that both average measures and change measures of theta can potentially each be uniquely informative about cognition. Indeed, there is increasing interest in moment-to-moment brain activity and within-individual variability in this, with some evidence indicating that measures of variability ought not to be dismissed as random error but could be informative about neural development (Garrett et al., 2013). For example, relations have been found between EEG signal variability and age (Angulo-Ruiz et al., 2021), as well as face recognition and reaction time (McIntosh et al., 2008) and visual and auditory processing (Lippe et al., 2009). There have thus been calls for research to further investigate variability in neural activity as they may provide useful insights into the functioning of the human brain that other neural measures might not (Garrett et al., 2013).

3.1.2.2.1 Regional specificity of theta

Despite significant research into theta measures, less is known about their regional specificity in children aged between 2 and 5 years. In some studies in this age range, theta power was specific to frontal brain regions (Canen & Brooker, 2017), whereas other work supports greater power in posterior areas or across widespread regions (Cuevas et al., 2012). It is possible that these varying findings are due to differences in participant age and experimental design.

In a sample of 2-year-olds, retrieval-related differences in theta power were found across widespread areas of the scalp, whilst encoding-related differences were specific to occipital regions (Cuevas et al., 2012). In a study involving two participant samples, (Orekhova et al., 2006) found that an increase in theta power during test compared to baseline conditions was predominantly frontal for infants but was more widespread for preschoolers (whose mean age was 5 years; range from 3 years, 8 months to 6 years, 11 months). Specifically, preschoolers' theta increases during exploratory behaviour were predominantly in anterior regions, whilst during social stimulation this was in posterior regions. In younger infants, the degree of difference in theta power during social and non-social conditions became more regionally widespread and prominent from 6 to 12 months of age (Jones et al., 2015), thus this pattern may continue into toddlerhood and preschool age leading to potentially changing patterns of theta power dominance. Existing research is limited on how these patterns may develop between 2 and 5 years of age, though it seems possible that there may be a shift from greater theta power in frontal to posterior regions in response to social stimuli during this period.

Though questions remain about the specific functioning and development of theta rhythms, there is substantial work indicating that it plays an important role in development (Adam et al., 2020; Kucewicz & Kamiński, 2022), thus supporting theta power as a potential candidate metric for work to optimise brain function. However, despite its prominent use in developmental neuroscientific research, reliability estimates of measures of theta power are lacking. One study which investigated test-retest reliability of EEG connectivity found that connectivity in the theta, along with lower alpha, band was particularly reliable compared to other frequency bands (van der Velde et al., 2019). Whilst this might indicate stability of theta connectivity over this 1-week period, this may not be true for theta power; additionally, this work was conducted in 10-month-old infants.

This means findings relating to theta power metrics currently ought to be taken with caution, as they may not be assessing a reliable measure of neural activity.

3.1.2.3 Conclusion

In conclusion, theta and alpha rhythms constitute a considerable portion of young children's EEG, though the distribution of dominance between them is a little unclear in the first few years of life. Measures of theta and alpha power have both been implicated in key cognitive processes, with evidence that alpha may be involved in control of attention and access of information, whilst theta may be particularly important for the processing of new information (Klimesch, 2012). Both are commonly used metrics which appear to relate to young children's cognition and may be good candidates for future research to develop interventions for brain optimisation; they would thus also benefit from further work investigating the reliability of measures for use in future work.

3.1.3 Relative versus absolute power

Though EEG power has thus far been discussed as a single concept, there are different ways of measuring EEG power, with analyses traditionally considering absolute or relative power in a particular frequency band. Absolute power refers to the absolute form of energy in a frequency band whilst relative power is energy in a frequency band divided by total energy from a broader frequency range (Govindan et al., 2017). Whilst absolute power is a measure of spectral power taken with minimal computations, additional calculations to minimise over or underestimation of EEG oscillations in a particular frequency band are performed to obtain relative power. EEG power shows an exponential decrease in relation to increasing frequencies, meaning measures of low frequency bands (i.e. theta) are much higher than in high frequency bands (i.e. alpha). As a result, small differences in power at higher frequencies may not be detected when only using absolute power. Absolute power may also be influenced by different EEG amplifiers (Kim, 2018) or inter-individual differences. In addition, absolute power is more sensitive to artifact activity than relative power (Panov, 2023), which may be particularly problematic in research with young children, where artifacts are likely more frequent. By creating a measure of power which is relative to a measure of overall power across wider frequencies, differences resulting from these factors can be controlled for, and power differences across frequencies are scaled to the same proportional measure.

Whilst relative power measures take into account differences in dominant frequency bands across different groups, they may also be sensitive to misinterpretation. In addition, there is no standardised method for calculating relative power, meaning comparison between studies is limited (Kim, 2018). A common method is to divide absolute power in the frequency band of interest by absolute power across a broader range of frequencies, though even when using this method, different studies have used a variety of frequency ranges for comparison. Work looking at relative theta/ alpha power has used ranges including 3-18Hz (Marshall et al., 2004), 2-45Hz (Hillard et al., 2013), 1-30Hz (Finnigan & Robertson, 2011; Nishiyori et al., 2021) and 1.5-22.5Hz (Somsen et al., 1997); a 2-20Hz range was chosen here in line with previous research (Markovska-Simoska & Pop-Jordanova, 2017) and to minimise influence from higher frequencies. Given that both absolute and relative power are used in developmental research, there are limitations with both, along with suggestions that the most reliable results can be ascertained by consideration of both absolute and relative EEG powers (Stroganova & Orekhova, 2007), the current study used both absolute and relative measures of theta and alpha power.

Some research suggests that relative power measures show higher reliability scores than absolute power (Fernández et al., 1993), whilst others indicate the opposite (Ma et al., 2019). Despite some work indicating that both absolute and relative theta power has poor reliability (Ma et al., 2019), other work has found fair to good reliability of relative theta and alpha power using a wireless EEG system at day, week and month long intervals (Rogers et al., 2016). Other work supports moderate to good test reliability of relative alpha power (Metzen et al., 2022) with some suggestion that absolute alpha had better reliability than relative alpha (Ma et al., 2019). In a large sample using the same mobile EEG system as the current study, good reliability was found for both absolute and relative theta and alpha power metrics in 12-month-old infants, though reliability measures for these frequencies were lower than for higher frequencies (Troller-Renfree et al., 2021). For absolute measures, a minimum of 20 1-second segments were needed to achieve good reliability, whilst relative measures required only 15 segments; excellent reliability was achieved with 40 and 35 trials respectively. In one lab-based study which included infants and pre-schoolers, good reliability was achieved for infants with a minimum of 10 1-second trials of EEG data, whilst this could be achieved from 5-year-olds with a minimum of 20 trials, depending on the method of artefact identification used (Leach et al., 2020). Of note, reliability levels might differ depending on a number of factors including time interval, age of participants, experimental task design and EEG system. Given these inconsistencies, it was difficult to form strong hypotheses about the levels of reliability for each measure in the current study, though fair to good reliability was expected in line with other developmental work (Popov et al., 2023; Troller-Renfree et al., 2021; van der Velde et al., 2019).

3.1.4 Current study

In the current study a test-retest design was used to assess the reliability of a wearable, wireless Enobio EEG system with typically developing 2.5- to 4-year-olds. Children watched a battery of visual tasks; the current chapter considers EEG measures whilst children were viewing social and non-social videos. It specifically focussed on alpha and theta power measures during video viewing, as these are thought to be involved with cognition and learning, and are increasingly used in developmental work (Begus & Bonawitz, 2020; Braithwaite et al., 2019; Klimesch, 2012; Orekhova et al., 2006). Several EEG power measures were used in this study, including average power over the whole video, average power in each of the first and second 30 seconds of video viewing, and average power in each one second segment of video viewing. Measures were calculated for each of the theta (3-6Hz) and alpha (6-10Hz) frequency bands.

For each measure, it was first investigated whether differences were found between frequency bands (alpha versus theta), across brain regions (frontal versus posterior), across video conditions (social versus non-social), and between multiple presentations of each video. Given work indicating increased alpha power in posterior regions during attention to visual stimuli (i.e Orekhova et al., 2001), greater average alpha power in posterior versus frontal regions was predicted. Since parietal and occipital alpha power have been related to memory encoding (Cuevas et al., 2012) and since social interaction is considered to provide individuals with a large amount of information, it seems possible that alpha power may be larger in the social compared to non-social video condition. For theta, given the lack of research using similar conditions with this specific age range, it was difficult to make clear predictions about regionally specificity, however it was expected that average theta power would be greater in the social compared to the non-social condition (Jones et al., 2015). By contrast, it was predicted there would be no difference in the degree of theta power change across conditions, with the caveat that this is based upon findings in much younger children (Jones et al., 2020). Though it is not exactly clear when the dominant frequency shifts from theta to alpha, it appears this may occur sometime between 10-months- and 7-years-old. Participants in the current study were within this age range, therefore comparing power, particularly relative power, between theta and alpha may be informative about the contributions of each frequency band at this age.

In addition to these comparisons, for measures of power change, it was also established whether any differences in power were found over the course of video viewing. For theta power, increases over the course of video viewing were expected, in line with previous infant work (Braithwaite et al., 2019; Jones et al., 2020), though this is, to my knowledge, the first time these types of analyses have been conducted in toddlers/ preschoolers. In line with previous work with preschoolers (Meyer et al., 2019), theta power was additionally considered across multiple viewings of the same video, with the prediction that theta would increase over subsequent viewings in line with increased task engagement. Given similar findings in this paper for alpha frequency bands, similar predictions were also made for alpha power change.

Following condition comparisons, assessments of test-retest reliability for each EEG power measure were conducted. As there is increasing interest in what variability might tell us about brain function (Garrett et al., 2013), the test-retest reliability of three measures of variability were additionally considered, in order to investigate whether how much EEG power varied over the course of video viewing was somewhat consistent within individuals. Given the individual differences in children's EEG, individuals' EEG responses were compared at two different time points, enabling us to determine how reliable the current measures were. In line with other work investigating the test-retest reliability of other EEG measures in young children, it was predicted that there would be reasonable-to-good levels of reliability for measures of power (Haartsen et al., 2021; van der Velde et al., 2019), though no specific hypotheses were made about which measures would be most reliable.

To summarise my hypotheses are as follows:

Alpha power

1. *Average alpha power will be greater during social than non-social condition*
2. *Posterior alpha power will be greater than frontal alpha power (i.e Orekhova et al., 2001)*

3. *Alpha power will increase over subsequent viewings of the same video (Meyer et al., 2019)*
4. *There will be reasonable test-retest reliability (of values between 0.56-0.76) for alpha power measures for each viewing of a video (Haartsen et al., 2021; van der Velde et al., 2019)*

Theta power

5. *Average theta power will be greater during social than non-social condition (Jones et al., 2015)*
6. *Theta power will increase over the course of video viewing (Braithwaite et al., 2019; Jones et al., 2020)*
7. *No difference in the degree of change across social and non-social videos (Jones et al., 2020)*
8. *Theta power will increase over subsequent viewings of the same video (Meyer et al., 2019)*
9. *There will be reasonable test-retest reliability (of values between 0.56-0.76) for theta power measures for each viewing of a video (Haartsen et al., 2021; van der Velde et al., 2019)*

3.2 METHODS

3.2.1 Participants

Participants were sixty-one (34 female) typically developing full-term children from the Greater London area. Participants were recruited from the Birkbeck Babylab database when they were between 30 and 48-months-old and were invited to attend two visits to the lab. Where possible, there was an interval of 1 to 2 weeks between visits, though sometimes scheduling difficulties meant they were further apart (see Table 3.1 for details). One to two weeks was chosen as the interval as shorter intervals might lead to repetition effects in the neural responses and data loss, whereas longer intervals may reflect developmental change rather than stability of the measures (Blasi et al., 2014; Haartsen et al., 2020).

At each visit, the study was explained to parents/ caregivers upon arrival at the Birkbeck Babylab and they then provided written informed consent. Participants received a certificate at the end of their first visit and a Birkbeck Babylab t-shirt or tote bag at the end of their second visit. Ethical approval was received for this study from the Department of Psychological Sciences ethics committee at Birkbeck, University of London (ref. No 171874).

Table 3.1 and Table 3.2 show demographic statistics for all recruited participants, for participants who provided usable data for all three videos from the test session ('test sample') and the retest session ('retest sample'). The test sample consists of 27 participants (12 female) and the retest sample consists of 27 participants (15 female).

Table 3.1: Demographic data for the recruited sample, test sample and retest sample. The test and retest samples were defined as children who provided any usable data for that session.

		N	*M*	*SD*	Min.	Max.
Recruited sample	Participant age at first visit (months)	61	38.36	4.66	30	49
	Participant age at first visit (days)	61	1181.49	143.30	936	1514
	Time between visits (days)	51	9.63	4.65	7	28
	Number of bedrooms in household	58	2.79	1.14	1	5
	People per bedroom	58	1.47	0.47	0.75	3
	Percent of English heard at home/ nursery	60	89.88	15.80	30	100
Test sample	Participant age at first visit (months)	44	38.09	5.09	30	49
	Participant age at first visit (days)	44	1173.30	155.38	936	1514
	Time between visits (days)	40	9.95	5.03	7	28

	Number of bedrooms in household	43	2.94	1.20	1	5
	People per bedroom	43	1.42	0.46	0.75	2.50
	Percent of English heard at home/ nursery	44	90.58	16.08	30	100
Retest sample	Participant age at first visit (months)	40	38.13	4.76	30	49
	Participant age at first visit (days)	40	1174.38	145.74	936	1514
	Time between visits (days)	40	9.80	4.99	7	28
	Number of bedrooms in household	39	2.87	1.20	1	5
	People per bedroom	39	1.41	0.43	0.75	2.5
	Percent of English heard at home/ nursery	40	89.95	17.18	30	100

Table 3.2: Demographic data for the recruited sample, test sample and retest sample. The test and retest samples were defined as children who provided any usable data for that session.

		Recruited sample		Test sample		Retest sample	
	Category	*N*	Percent	*N*	Percent	*N*	Percent
Annual household income	Less than £20,000	4	7%	2	4.8%	1	2.6%
	£20,000 - £29,999	5	8.8%	5	11.9%	4	20.5%
	£30,000 - £39,999	5	8.8%	3	7.1%	2	5.3%
	£40,000 - £59,999	7	12.3%	6	14.3%	5	13.2%
	£60,000 - £79,999	6	10.5%	4	9.5%	5	13.2%
	£80,000 - £99,999	12	21.1%	11	26.2%	10	26.3%
	£100,000 - £149,999	14	24.6%	8	19%	8	21.1%
	More than £149,999	4	7%	3	7.1%	3	7.9%
	Missing	4	NA	2	NA	2	NA
Medical history	Child required speech/ language therapy	4	6.7%	4	9.1%	1	2.5%
	Family member required speech/ language therapy	2	3.3%	2	4.5%	2	5.0%
	Missing	1	NA	0	NA	0	NA

3.2.1.1 Data inclusions

Of the 61 participants, nine participants provided no usable EEG data; two due to technical issues and seven who refused the EEG cap at the first visit. Of the 52 participants from whom data was recorded at the test session, seven participants provided no usable EEG data from their second visit: five participants did not return for a second visit due to scheduling difficulties/ logistics, one did not return due to parental concerns and one child who had provided data at the first visit refused to wear the EEG cap at the second visit (see Figure 3.1). Data processing was done on all data provided from 52 participants in the test session and 45 participants in the retest session; due to equipment failures, not all EEG files contained usable data for each video at each session. Videos were split into three repetitions of both social and non-social conditions at test and retest sessions, meaning the maximum number of trials a participant could provide at each session was six. Table 3.3 shows the number of participants providing data for no, some and all conditions for test, retest and both sessions; Table 3.4 shows statistics for the number of conditions provided by children providing some data for test and retest sessions. All usable data provided by participants will be included in analyses as far as possible.

Figure 3.1: Exclusion information for the recruited sample

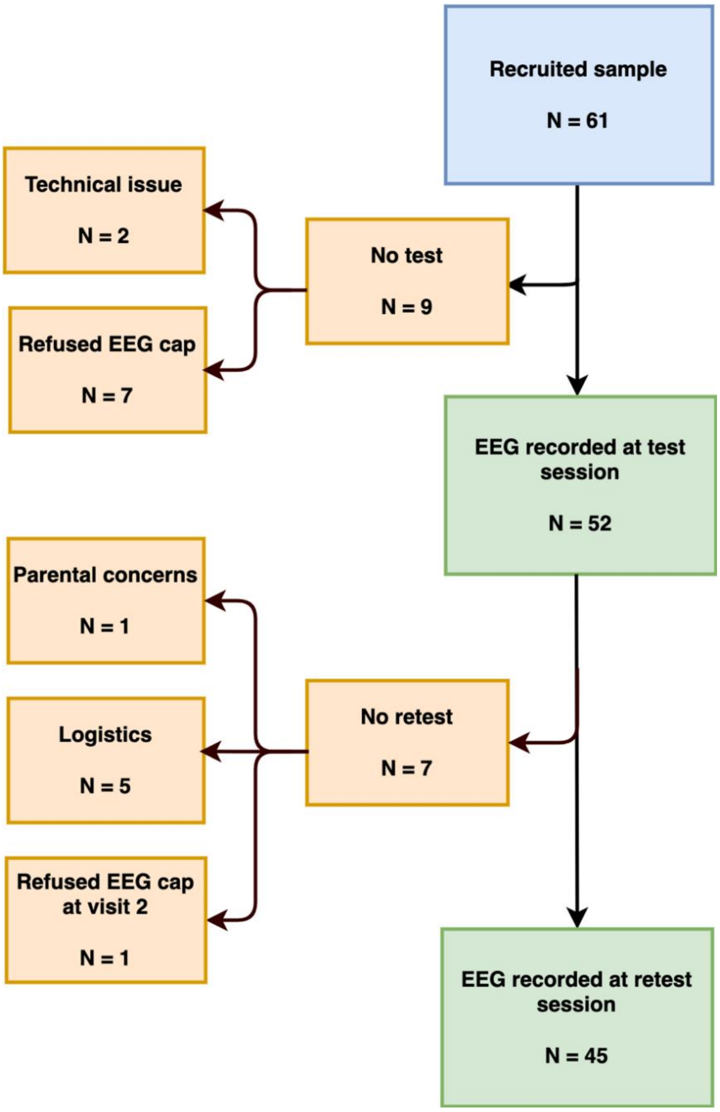

Recruited sample
N = 61
Technical issue
N = 2
No test
N = 9
Refused EEG cap
N = 7
EEG recorded at test session
N = 52
Parental concerns
N = 1
No retest
N = 7
Logistics
N = 5
Refused EEG cap at visit 2
N = 1
EEG recorded at retest session
N = 45

Table 3.3: Number of participants who provided data for none, some and all video conditions and repetitions.

	Test	Retest	Both
No data	8	5	3
Data for some conditions missing	17	13	13
Data for all six conditions	27	27	16

Table 3.4: Descriptive statistics showing the mean and range of the number of trials provided for each of the test and retest sessions for participants who had data for some conditions missing

	N	Mean	SD	Min.	Max.
Test	17	3.71	1.45	1	5
Retest	13	4.31	0.86	2	5

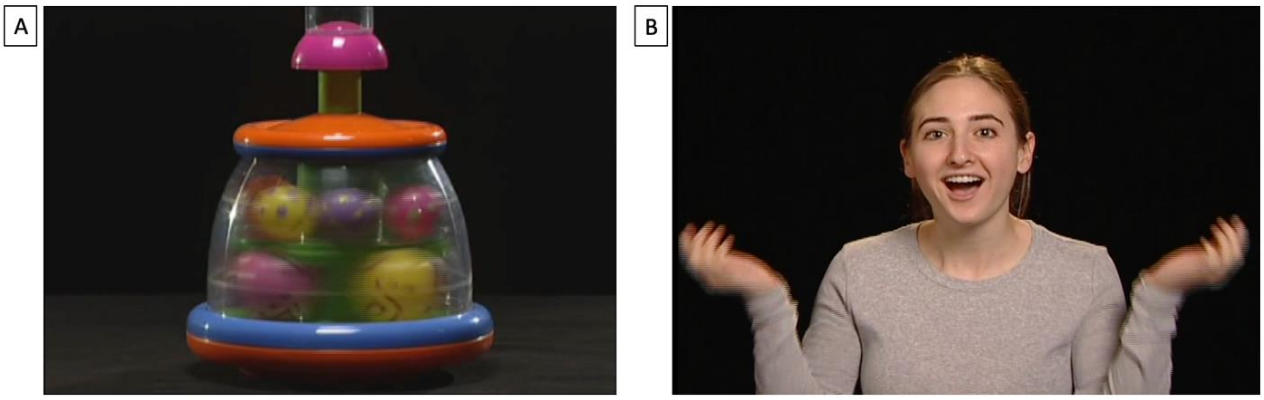
3.2.2 Materials and stimuli

3.2.2.1 EEG session

At each visit to the lab, EEG data was gathered during a battery of tasks. The battery consisted of both visual and auditory tasks designed to be interesting and engaging. The length of the whole battery was approximately 35 minutes, whilst the length of the visual tasks ranged from 22-27 minutes depending on participants engagement. The auditory task was an auditory oddball task. Visual tasks included fast event-related potentials (ERPs) (see Haartsen et al., 2021) interspersed with short cartoon clips and dynamic videos. Dynamic videos are the focus of this chapter.

Dynamic videos were (a) non-social: clips of child-appropriate dynamic toys moving with no visible human action (i.e. a ball falling through a stacking ball drop toy) (Figure 3.2a) and (b) social: short clips of women singing nursery rhymes with actions and sound (Figure 3.2b). Social and non-social videos were 60 seconds long and began with a gaze-contingent fixation stimulus (Flaticon image, 3cm by 3cm); once participants fixated the fixation stimulus, it was replaced by the video. Participants saw both the social and non-social videos three times each.

Figure 3.2: Screenshots from (A) non-social and (B) social videos



3.2.2.1.1 Stimuli presentation

Stimuli were presented on an external monitor (Asus VG248, 24-inch screen size, 1902 x 1080 resolution at 60Hz) via a MacBook Pro (15-inch, part number MR942B/A, with an eight-generation Intel i7 6-core processor, 2016). A portable Tobii Pro X2 eye-tracker was attached to the bottom of the participant’s monito and connected to the testing MacBook Pro via a wire. Eye-tracking data were recorded with Tobii Pro SDK 3Manager, whilst visual stimuli were presented and data were saved using the stimulus presentation framework, Task Engine (https://sites.google.com/site/taskenginedoc/; Jones et al., 2019) which is optimised for standardised EEG and eye-tracking data collection. The framework was run on a macOS High Sierra 10.13.6 system, in MATLAB R2017a, with Psychtoolbox 3.0.14, Gstreamer 1.14.2 for stimulus presentation, and a Lab Streaming Layer to connect to the EEG system; EEG and ET task events were time stamped and saved in MATLAB format. Sessions were also recorded using a webcam (HD Pro Webcam C920) attached to the top of the participant’s monitor and the Open Broadcaster Software (OBS) on the MacBook Pro.

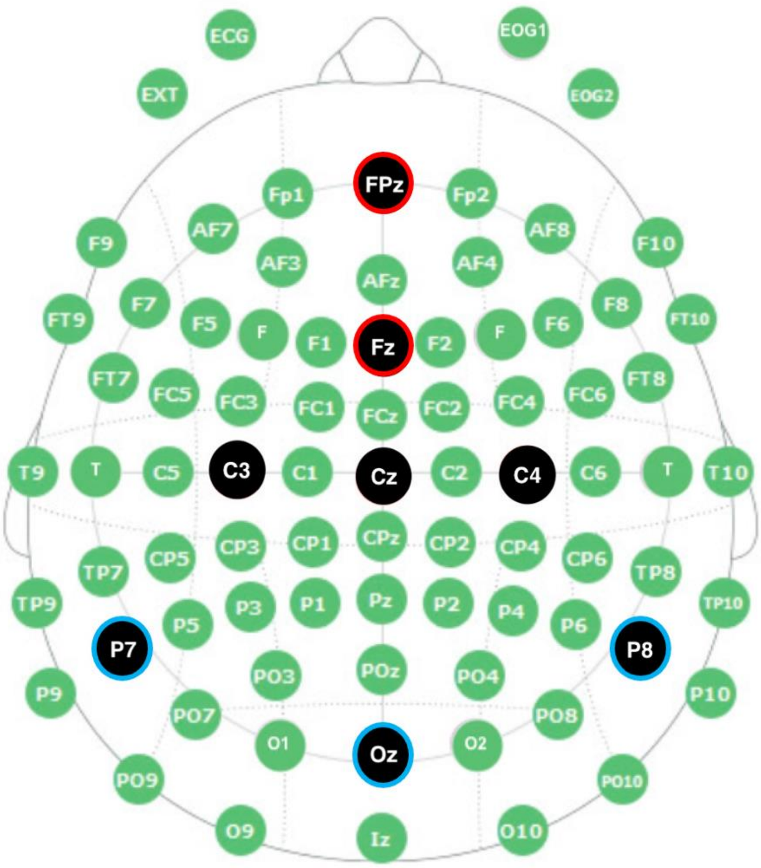
3.2.2.1.2 EEG recording

During stimuli presentation, EEG was simultaneously recorded using a wireless geltrode Enobio EEG system (NE, Neuroelectrics, Barcelona, Spain). Data was recorded from eight electrodes placed at FPz, Fz, Cz, Oz, C3, C4, P7 and P8 (see Figure 3.3 for channel layout) and transferred to the MacBook Pro via a Bluetooth connection. A low-density array was used to facilitate good quality data collection by balancing this with quantity of channels. Young children typically have a considerable amount of hair, meaning some time is required to seat electrodes and ensure data quality, however they also have limited patience and often low tolerance during application, meaning application time needs to be kept to a minimum. Using large arrays can lead to greater attrition rates, as some children's patience may be crested and they may end up providing no data, though they may have provided data with a smaller array.

For most participants the CMS and DRL reference electrodes were attached to an ear clip placed on one of the participants' ears. Eight children did not tolerate the earclip; for these participants the CMS and DRL were instead attached to sticktrodes placed on each mastoid. Before applying either the ear clip or the sticktrodes, the area was gently cleaned with antibacterial spray. The software Neuroelectrics NIC 2.0 (Barcelona, Spain) was used to record data with a sampling rate of 500Hz and to visually monitor data quality throughout the session. The Neuroelectrics Quality Index (QI) calculates line noise (power in the range of 49 - 51Hz), main noise (power in the range of 1-40Hz) and the offset of the signal every 2 seconds. NIC software displays a colour code indicating the Quality Index for each recorded electrode; green indicates low QI and good data quality, orange indicates average data quality and red indicates high QI and poor data quality.

This set up is also described by Haartsen et al. (2021) that paper also includes data from the fast ERPs collected as part of this study.

Figure 3.3: Layout for Enobio cap. Channels used in the Braintools paradigm are marked in black.



3.2.2.2 Questionnaires

Parents/ caregivers filled out the following questionnaires as part of this study: *a demographic questionnaire, medical questionnaire, language questionnaire and a parental feedback form.*

3.2.3 Procedure

Exactly the same procedure was followed at both visits. At each visit there was always one main researcher (EB or TDB) along with at least one helper. After gathering written informed consent from a parent/ guardian and allowing the child some time to settle into the lab environment, the experiment began. The experiment consisted of an EEG session during which simultaneous EEG and eye-tracking data were collected whilst children watched a battery of tasks, and a number of questionnaires were completed by a parent/ guardian.

3.2.4 EEG Session

During the experiment, children were seated on their parent or caregiver's knee approximately 60cm away from the participant monitor. During EEG cap application, participants were presented with cartoon videos (Disney or Pixar clips) and entertained by one researcher, whilst another applied the EEG cap. Once the EEG cap had been fitted, experimenters moved out of view behind a screen, from which stimuli and testing was controlled via a MacBook Pro. If participants became restless and distracted during the battery, one experimenter sometimes moved to sit next to them to encourage them to continue watching.

3.2.4.1.1 EEG cap application

The EEG cap had been prepared by a researcher prior to the participant's arrival so that it contained geltrodes in each of the eight electrode locations used in this study. The cap (with geltrodes inserted) was placed onto the participant's head and the chin strap fastened. The researcher then inserted gel to each of the eight geltrodes, before attaching electrodes to each. The earlobe (or mastoids where the earclip was refused) was cleaned and references (either earclip or mastoid sticktrodes) were applied. The electrode bundle was then plugged in to the NIC box and the programme NIC2.0 was used to visually assess the EEG signal. Once the EEG signal was deemed good enough, eye-tracking calibration began.

3.2.4.1.2 Eye-tracking calibration

Eye-tracking calibration was performed by presenting a coloured spiral shape which moved between five locations of the screen: the centre and each of the four corners. As it did so, children's eye-gaze was detected and noted whether this was valid or not. Researchers could choose to rerun the calibrations until it became valid. Researchers did not proceed from calibration until at least four of the five locations on the screen were valid. Eye-tracking data was gathered throughout EEG recording. Social and non-social videos began when participants looked at an attention-grabber in the centre of the screen. Visual attention grabbers were small icons of everyday items (i.e. a cupcake, a bowl, etc.) inside a circle in the centre of the screen. Sometimes there were technical difficulties or children were moving around too much for the eye-tracker to pick up their gaze. In these cases, the experimenter used the video recording of the child to assess when they were looking at the stimulus on the screen and manually started the videos by pressing the ‘tab’ key. Parents/ caregivers were asked to wear a pair of plastic shutter glasses to ensure the eye-tracker picked up only the child’s eye gaze. After a successful eye-tracking calibration, the stimulus battery began.

At the first visit parents/ caregivers filled out a demographic questionnaire, a medical questionnaire, and a language questionnaire; at the end of the second visit they filled out a feedback form asking their thoughts about the study.

3.2.5 Data processing

3.2.5.1 EEG

Data recorded during the social and non-social videos were cut from the rest of the recording and then pre-processing began. Data were pre-processed using a combination of in-house written scripts and Fieldtrip MATLAB scripts (Oostenveld et al., 2011) in MATLAB R2021a. First, data validity was checked and corrected where necessary and possible, e.g. cases with technical issues during data collection or saving, or inaccurate timing or absence of EEG markers. Enobio data were converted into Fieldtrip format for further pre-processing.

Data were split into separate files based on whether they were the first, second or third occurrence of each video; for each file continuous data were then segmented into 1-second epochs with no overlap. Data were detrended and a 0.1-48 Hz bandpass filter was applied to filter out 50Hz line noise and higher-frequency noise from muscle artifacts. Artifacts were identified using both automatic and manual detection. In automatic detection, trials were marked bad if the signal exceeded thresholds -150 to 150 µV, or if a flat signal or a jump of greater than 20 µV were detected; these trials were then removed. In manual detection, trials were removed if more than one channel displayed artifacts or if there were artifacts in the signal from any channels in the particular regions of interest (i.e. Fz, Fpz, P7, P8 or Oz). Manual artifact detection was done by trained volunteer student; additionally, one researcher checked all files to ensure consistency across students.

Artifact-free data were then detrended and re-referenced on a trial-by-trial basis to Cz, or the average of C3 and C4 if Cz contained artifacts. If C3 or C4 also contained artifacts, the whole trial was excluded from further analysis. As there is currently very little work using only 8-channels, this reference method was chosen in line with other work using the same low-density array (Haartsen et al., 2021; Throm et al., 2023). Previous work that used trial-by-trial referencing found that this led to greater detention and did not significantly impact results (Del Bianco et al., 2024). To assess whether reference location impacted measures, additional analyses were conducted in which reference location (either Cz or the average of C3 and C4) was used as a covariate. These results are reported in appendix B. Following re-referencing, a fast-Fourier transformation was applied to re-referenced data to ascertain power (μv^2) per electrode in 1Hz bins for 1-48Hz. Power data were split into social and non-social conditions, and absolute and relative power were calculated. Mean power over the theta (3-6Hz) and alpha (6-10Hz) frequency bands were used as absolute power. Relative power was calculated by dividing the sum of power in the theta and alpha frequency bands by the sum of power in the 2-20Hz range, as in Segalowitz et al. (2010).

3.2.5.1.1 EEG measures

Several measures of power were used; for each of the theta and alpha frequency bands both absolute and relative power measures were calculated for each of the frontal and posterior regions. Measures in the frontal regions were taken as the average power over channels Fpz and Fz (circled in red in Figure 3.3), whilst posterior regions averaged power over P7, P8 and Oz (circled in blue in Figure 3.3). Overall average power was calculated as mean power over the whole 60 seconds of video viewing, for any 1-second segments that were left after processing and artifact removal. Mean EEG power in the first and second 30 seconds of video viewing were also calculated, such that change in power could be investigated. In initial condition analyses, ‘half’ was used as a within-subjects variable with two levels (first and second half of the video) to assess whether mean power differed at the start and end of the video. For later reliability analyses an index was calculated as the difference between the two halves (mean in first half minus mean in second half); this was then used in assessments of reliability.

A more continuous measure of power over the course of video viewing was also calculated, using average power in each 1-second segment of video viewing to consider power change over segment number, in addition to other condition comparisons. Linear mixed models were fitted using power as a dependent variable and segment number as a fixed effect (more details in sections 3.3.1.3 and 3.3.2.3 below). An intercept value was also obtained from this model for each subject, which was then used as a second measure of individual average theta power.

To understand whether individual variability in EEG power over the course of video viewing was reliable between test and retest sessions, residual standard error, standard deviation, and standard error of intercept estimates were also considered. These measures were chosen to include measures of variability around mean power and around regression lines (i.e. in relation to power change).

3.2.5.2 Questionnaires

The demographic, medical and language questionnaires were used to determine exclusions and to provide demographic information for participants in the current study.

3.2.6 Statistical analyses

For average power metrics, linear models were fitted with average power over the whole 60 seconds of video viewing as the dependent variable and independent variables of EEG signal (alpha versus theta), region (frontal or posterior), video number (first, second or third) and video condition (social versus non-social). To compare power across the two halves of video viewing, linear models were conducted using the same independent variables, with video half (first or second) additionally included and average power across each half of video viewing as the dependent variable. ANOVAs were run on the linear models as they enable more precise and powerful analyses compared to t-tests.

For continuous power over the course of a video viewing, a linear mixed model was used with power per 1 second segment as a dependent variable and segment number as a fixed effect. Session (test or retest), EEG signal (alpha or theta), region (frontal or posterior), video number (first, second or third) and video condition were additionally included as fixed effects, whilst participant ID was included as a random effect with variable intercepts.

Analyses were repeated for each of the test and retest sessions, and for both absolute and relative EEG measures. Children were only included in each of the analyses if they had sufficient artifact-free EEG data for each analysis. For average power measures and continuous power change, any videos which had fewer than 10 seconds of useable EEG were excluded from further analyses. For average power in each half, videos were excluded if either of the first or second 30 seconds of the video had fewer than 5 seconds of usable EEG. Assumptions were checked before models were fit: any concerns about these were reported in the relevant section in results.

Intraclass correlations (ICC) were then performed to assess the test-retest reliability between the test and retest sessions (as in Haartsen et al., 2021; van der Velde et al., 2019) of the following metrics:

1. Average power over the whole 60 seconds of video viewing
2. Difference between average power in the first 30 seconds of video viewing and in the second
3. Intercepts provided by the linear mixed model
4. Variability measures

I used an ICC model of type (3,1) (two-way mixed effects, single measurement, absolute agreement) as used in other similar designs (Byun et al., 2016; Heilicher et al., 2022) and as recommended by Koo and Li (2016). ICC calculations were conducted in MATLAB R2021a using the 'A-1' specification in the "ICC" function (Salarian, 2023), this corresponds to the 'ICC(A,1)' specification in McGraw and Wong (1996) notation. The formula for this ICC is:

$$\frac{\mathrm{MS_R} - \mathrm{MS_E}}{\mathrm{MS_R} + (k-1)\mathrm{MS_E} + \frac{k}{n}(\mathrm{MS_C} - \mathrm{MS_E})}$$

MS_R is the variance between objects, MS_E is the error variability (i.e. mean squared error), MS_C is the variance between raters, k is the number of measurements per participant and n is the sample size. In this case, MS_R is the variance between participants, MS_C is the variance between sessions (test and retest), while k is two. ICC values usually range between 0-1, though negative values are also possible. Values less than 0 indicated a poor fit of the ICC, values close to 0 indicate poor test-retest reliability, whilst those close to 1 indicate excellent reliability. As is convention, ICC values were interpreted as follows: below 0.40 was poor; from 0.40 through 0.59 was fair, from 0.60 to 0.75 was good, and above 0.75 was excellent (as in Haartsen et al., 2021; van der Velde et al., 2019).

Most statistical analyses were conducted in R Studio (Core Team, 2023), though some simple statistics were conducted using MATLAB R2021a.

3.3 RESULTS

Analyses are first presented for absolute EEG power (section 1.4.1) followed by relative EEG power (section 1.4.2). Table 3.5 shows descriptive statistics for the number of trials for the whole and each half of videos in both test and retest sessions from all participants who provided any usable data. Figure 3.4 shows a visualisation of results for condition comparisons of average EEG power over the whole video, whilst Figure 3.5 shows these for analyses of average power in the first and second half of video viewing.

Table 3.5: Mean, standard deviation, maximum and minimum number of trials per video and condition during test and retest sessions

	TEST				RETEST			
	M	*SD*	Min.	Max.	*M*	*SD*	Min.	Max.
				First Video				
Social	47.49	9.36	20	63	46.73	11.42	19	64
Non-social	39.92	8.31	20	53	39.53	9.05	16	59
Social first half	21.16	5.59	5	30	21.27	5.85	7	30
Social second half	22.18	5.10	2	29	22.08	5.61	9	30
Non-social first half	20.18	4.83	11	28	19.95	6.08	5	29
Non-social second half	19.51	4.50	8	25	19.34	4.30	8	29
				Second Video				
Social	50.37	10.05	32	65	48.27	10.96	17	65
Non-social	42.74	12.78	14	61	43.86	10.93	14	59
Social first half	23.13	5.35	12	30	22.65	5.58	4	30
Social second half	23.21	5.60	11	30	22.49	5.32	8	30

Non-social first half	22.18	6.27	7	30	22.25	6.49	6	30
Non-social second half	20.21	7.24	3	30	21.14	5.99	6	28
				Third Video				
Social	50.40	12.16	17	64	48.20	13.34	15	63
Non-social	45.11	14.28	14	94	44.11	12.24	10	60
Social first half	23.60	5.95	7	30	21.91	6.79	6	30
Social second half	22.98	6.56	3	30	22.43	6.43	6	30
Non-social first half	23.00	7.99	3	52	21.53	6.87	2	29
Non-social second half	21.83	7.06	8	41	22.86	5.53	7	30

Figure 3.4: Heatmap of p-values for condition comparisons of average power over the whole video for absolute and relative power in test and retest sessions (as reported in sections 3.3.1.1 and 3.3.2.1)

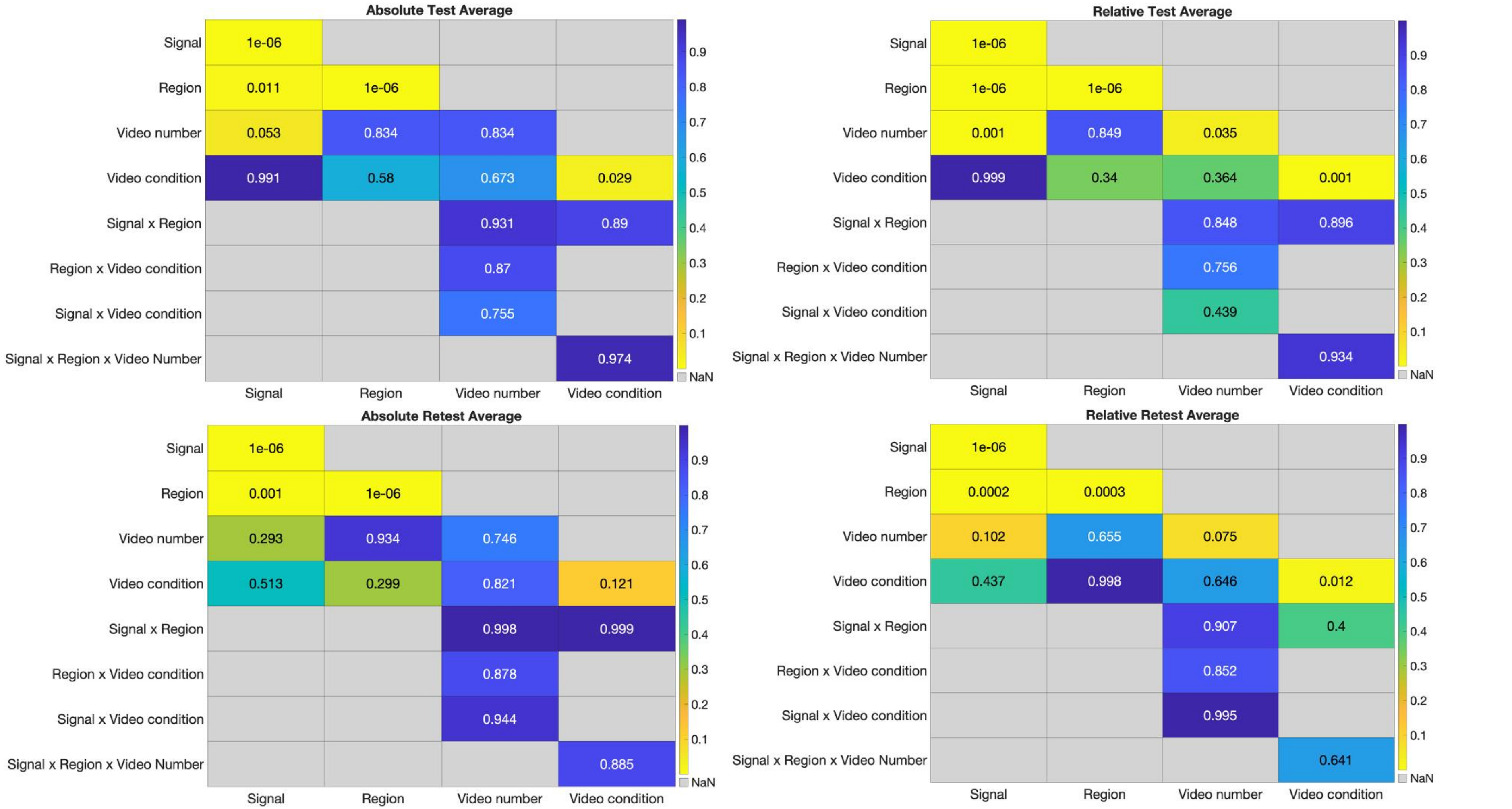
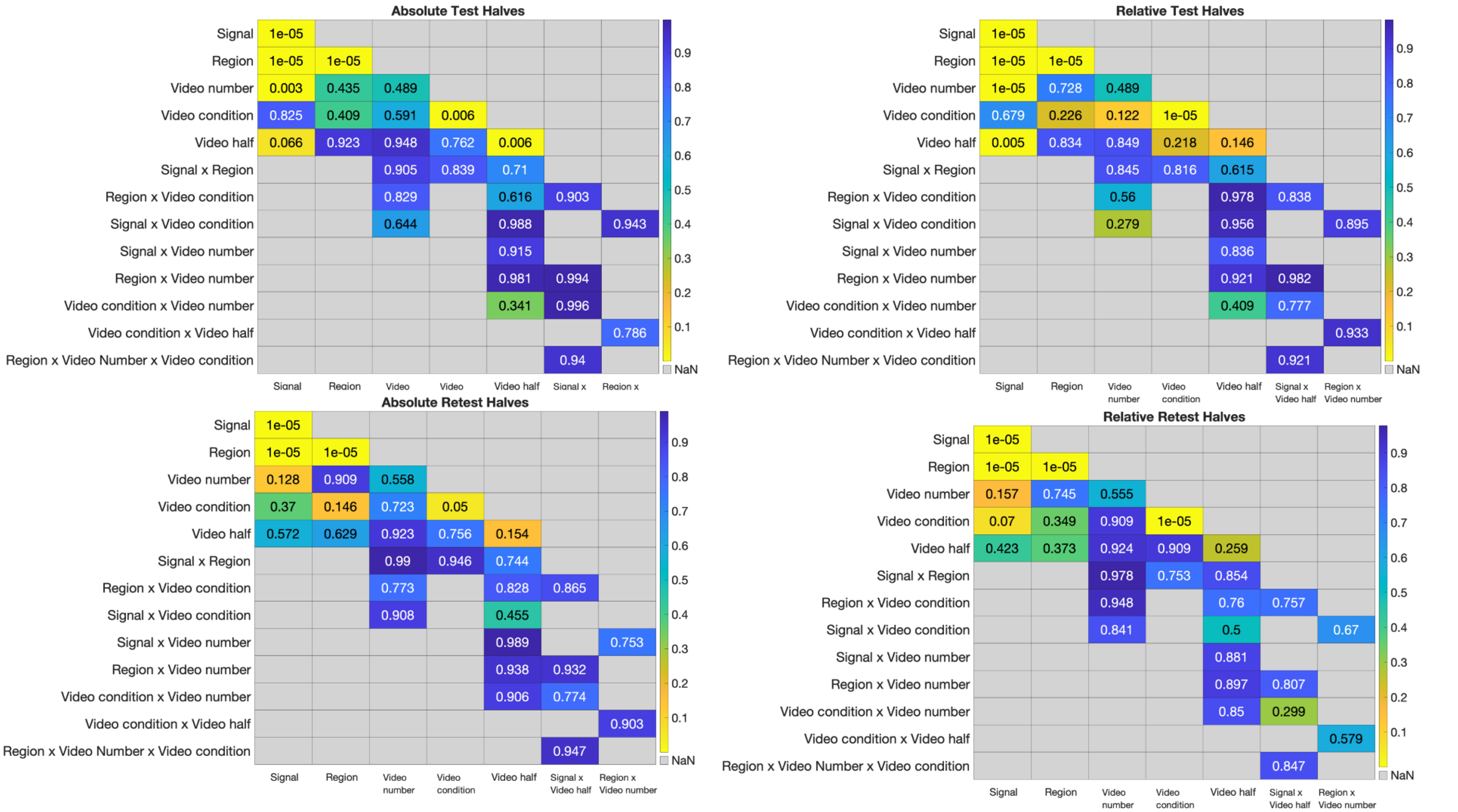


Figure 3.5: Heatmap of p-values for condition comparisons of average power in the first and second half of video viewing for absolute and relative power in test and retest sessions (as reported in sections 3.3.1.2 and 3.3.2.2)



3.3.1 Absolute power

This section contains condition analyses for each measure of alpha and theta power, namely average power over the whole video, average power in the first and second 30 seconds of video viewing and average power per second segment of video viewing. Some participants did not have EEG for both test and retest sessions, for all three videos, or for both social and non-social conditions. To maximise upon the available data linear regressions and linear mixed models were used to compare between conditions, as these do not require each participant to have data for every one condition. Descriptive statistics are also presented for all participants who provided EEG data for each condition.

3.3.1.1 Average power

To compare how average power differed across conditions within each of the test and retest sessions, repeated-measures linear models were fitted using the *lm* function in R (Core Team, 2023). Average power over the whole 60 seconds of video viewing was the dependent variable and independent variables of EEG signal (alpha versus theta), region (frontal or posterior), video number (first, second or third) and video condition (social versus non-social). Videos which had fewer than ten usable segments of EEG data were excluded from analyses and assumptions were checked. Analogous analyses were conducted using a minimum of 30 usable EEG segments per condition to check the robustness of effects; this did not change the pattern of findings (see appendix B1a for statistics). Identical models were fitted for data from the test and retest session; where significant interactions were found, follow-up pairwise contrasts were conducted using the *emmeans* function in R (Lenth, 2023) with a Bonferroni correction for p-values.

3.3.1.1.1 Test session

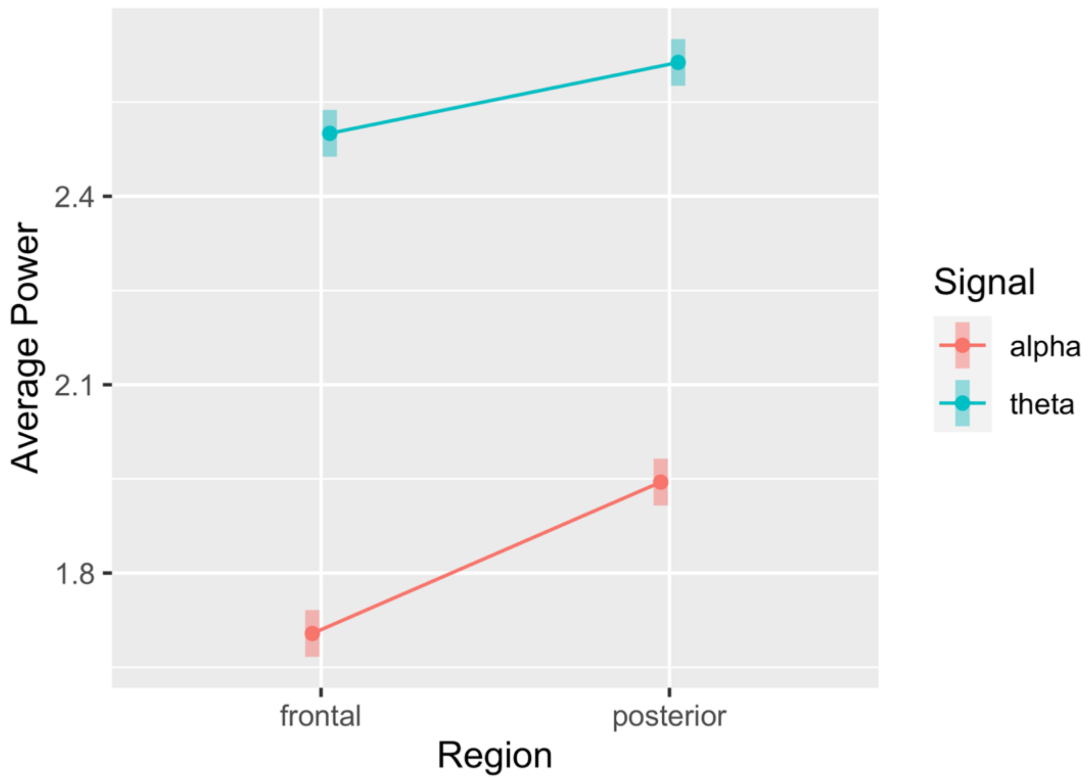
Table 3.6 show statistics from an ANOVA run on the linear model fit with data from the test session; Table 3.7 shows corresponding descriptive statistics. Significant results were found for the main effects of signal and region, and for the interaction between signal and region.

Means indicate that average theta power (M = 2.55, SE = 0.02, CIs [2.52, 2.59]) was higher than average alpha power (M = 1.82, SE = 0.02, CIs [1.79, 1.86]), and average power was higher in the posterior (M = 2.27, SE = 0.02, CIs [2.24, 2.31]) compared to frontal region (M = 2.10, SE = 0.02, CIs [2.07, 2.14]). Average power in the social condition (M = 2.22, SE = 0.02, CIs [2.18, 2.25]) was higher than in the non-social condition (M = 2.16, SE = 0.02, CIs [2.13, 2.20]).

Signal x region

Contrasts indicated that average theta power in the frontal (M = 2.50, SE = 0.03, [2.45, 2.55]) was significantly higher compared to posterior regions (M = 2.61, SE = 0.03, [2.56, 2.66]) [$t(900) = -3.03$, SE = 0.04, $p = .014$], and average alpha power was lower in the frontal (M = 1.71, SE = 0.03, [1.66, 1.75]) compared to posterior (mean = 1.94, SE = 0.03, [1.89, 1.99]) region [$t(900) = -6.61$, SE = 0.04, $p < .001$]; Figure 3.6. Posterior alpha power was significantly lower than posterior theta power, [$t(900) = -18.58$, SE = 0.04, $p < .0001$], and frontal alpha power was significantly lower than frontal theta power, [$t(900) = -22.16$, SE = 0.04, $p < .0001$].

Figure 3.6: Mean and confidence intervals for absolute frontal and posterior alpha and theta power over the whole video in the test session



3.3.1.1.2 Retest session

Similar to in the test session, significant main effects of signal and region were found, as well as a significant interaction between signal and region, however the main effect of video condition was not significant. Table 3.6 show statistics from an ANOVA run on the linear model fit with data from the retest session; Table 3.7 shows corresponding descriptive statistics.

Means indicate that average theta power (M = 2.62, SE = 0.02, [2.59, 2.65]) was higher than average alpha power (M = 1.80, SE = 0.02, [1.76, 1.83]), and average power was higher in the posterior (M = 2.33, SE = 0.02, [2.30, 2.37]) compared to the frontal region (M = 2.08, SE = 0.02, [2.05, 2.12]).

Signal x region

Contrasts indicated that average theta power in the frontal (M = 2.53, SE = 0.03, [2.48, 2.58]) was significantly higher compared to posterior regions (M = 2.71, SE = 0.03, [2.66, 2.76]) [t(840) = -4.91, SE = 0.04, p < .0001], and average alpha power was lower in the frontal (M = 1.63, SE = 0.03, [1.58, 1.68]) compared to posterior (M = 1.96, SE = 0.03, [1.91, 2.01]) region [t(840) = -9.46, SE = 0.04, p < .001]. Posterior alpha power was significantly lower than posterior theta power, [t(840) = -21.08, SE = 0.04, p < .0001], and frontal alpha power was significantly lower than frontal theta power, [t(840) = - 25.64, SE = 0.04, p < .0001]; Figure 3.7.

Figure 3.7: Mean and confidence intervals for absolute frontal and posterior alpha and theta power over the whole video in the retest session

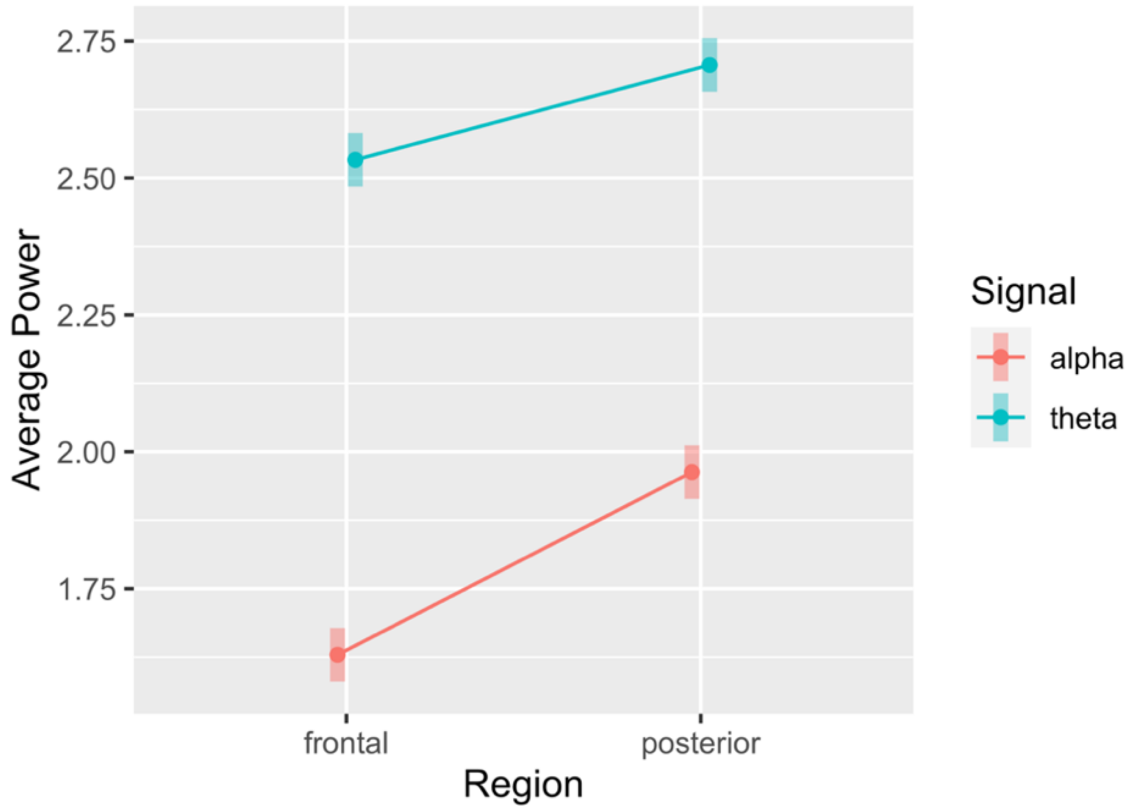


Table 3.6: F-value, p-value and partial eta squared effect size for ANOVAs on average absolute power over the whole video for test and retest sessions. Significant effects are indicated by *

	TEST			RETEST		
	F	*p*	η^2 (partial)	*F*	*p*	η^2 (partial)
Signal	831.73	<.001*	0.48	1089.77	<.001*	0.56
Region	46.57	<.001*	0.05	103.18	<.001*	0.11
Video number	0.18	0.834	0.00	0.29	0.746	0.00
Video condition	4.78	0.029*	0.00	2.41	0.121	0.00
Video condition x Video number	0.40	0.673	0.00	0.20	0.821	0.00
Signal x Region	6.41	0.011*	0.00	10.37	0.001*	0.01
Signal x Video number	2.96	0.053	0.00	1.23	0.293	0.00
Signal x Video condition	0.0001	0.991	0.00	0.43	0.513	0.00
Region x Video number	0.18	0.834	0.00	0.07	0.934	0.00
Region x Video condition	0.31	0.580	0.00	1.08	0.299	0.00
Region x Video condition x Video number	0.14	0.870	0.00	0.13	0.878	0.00
Signal x Video condition x Video number	0.28	0.755	0.00	0.06	0.944	0.00
Signal x Region x Video number	0.07	0.931	0.00	0.002	0.998	0.00
Signal x Region x Video condition	0.02	0.890	0.00	<0.001	0.999	0.00
Signal x Region x Video number x Video condition	0.03	0.974	0.00	0.12	0.885	0.00

Table 3.7: Mean, standard deviation and N for average absolute power over the whole video for test and retest sessions

Signal	*Region*	*Video number*	*Video condition*	TEST			RETEST		
				M	*SD*	*N*	*M*	*SD*	*N*
Theta	Frontal	First	Social	2.49	0.28	37	2.53	0.40	37
			Non-social	2.45	0.28	39	2.52	0.33	37
		Second	Social	2.52	0.34	38	2.57	0.38	37
			Non-social	2.48	0.33	39	2.51	0.36	36
		Third	Social	2.54	0.30	40	2.55	0.41	34
			Non-social	2.52	0.34	38	2.53	0.35	35
	Posterior	First	Social	2.60	0.42	37	2.75	0.34	37
			Non-social	2.54	0.39	39	2.62	0.34	37
		Second	Social	2.65	0.43	38	2.75	0.37	37
			Non-social	2.60	0.38	39	2.68	0.35	36
		Third	Social	2.68	0.41	40	2.74	0.34	34
			Non-social	2.57	0.40	38	2.70	0.32	35
Alpha	Frontal	First	Social	1.77	0.41	37	1.68	0.46	37
			Non-social	1.78	0.37	39	1.66	0.40	37
		Second	Social	1.71	0.38	38	1.63	0.37	37
			Non-social	1.62	0.30	39	1.64	0.39	36
		Third	Social	1.70	0.40	40	1.58	0.46	34
			Non-social	1.65	0.36	38	1.60	0.45	35

Posterior	First	Social	1.98	0.45	37	2.03	0.33	37
		Non-social	1.98	0.41	39	1.96	0.30	37
	Second	Social	1.97	0.51	38	1.98	0.34	37
		Non-social	1.87	0.43	39	1.95	0.30	36
	Third	Social	1.97	0.44	40	1.96	0.34	34
		Non-social	1.88	0.39	38	1.91	0.32	35

3.3.1.2 Power change: halves

To compare how average power differed over the course of video viewing, repeated-measures linear models were fitted using the *lm* function in R (Core Team, 2023) using average power in each of the first and second 30 seconds of video viewing as the dependent variable. Independent variables were EEG signal (alpha versus theta), region (frontal or posterior), video number (first, second or third), video condition (social versus non-social) and video half (first or second). Videos were excluded from analyses if either half had fewer than five usable segments of EEG and assumptions were checked. Robustness checks including a minimum of 15 usable EEG segments per half did not find a significantly different pattern of results (see appendix B2a). Identical models were fitted for data from the test and retest session; where significant interactions were found, follow-up pairwise contrasts were conducted using the *emmeans* function in R (Fox & Weisberg, 2019) with a Bonferroni correction for p-values.

3.3.1.2.1 Test session

A significant main effect of half was found, with means indicating that average power was higher in the second half (M = 2.22, SE = 0.01, CIs [2.19, 2.24]) compared to the first half (M = 2.16, SE = 0.01, CIs [2.14, 2.19]) of video viewing. All statistics from the linear model are reported in Table 3.8; descriptive statistics are in Table 3.9.

3.3.1.2.2 *Retest session*

In contrast to the test session, the main effect of video half was not found to be significant; nor were any interactions involving video half (see Table 3.8 and Table 3.9).

Table 3.8: F-value, p-value and partial eta squared effect size for ANOVAs comparing condition effects for average absolute power over the first and second half of video viewing for test and retest sessions. Significant effects are indicated by*

	TEST			RETEST		
	F	*p*	η^2 (partial)	*F*	*p*	η^2 (partial)
Signal	1485.77	<.0001*	0.46	1968.42	<.0001*	0.54
Region	86.92	<.0001*	0.05	186.64	<.0001*	0.10
Video number	0.72	0.489	0.00	0.58	0.558	0.00
Video condition	7.54	0.006*	0.00	3.85	0.050	0.00
Half	7.56	0.006*	0.00	2.036	0.154	0.00
Signal x Region	5.73	<.001*	0.00	18.81	<.001*	0.01
Signal x Video number	1.85	0.003*	0.00	2.06	0.128	0.00
Region x Video number	0.83	0.435	0.00	0.10	0.909	0.00
Signal x Video condition	0.05	0.825	0.00	0.80	0.370	0.00
Region x Video condition	0.68	0.409	0.00	2.11	0.146	0.00
Video number x Video condition	0.53	0.591	0.00	0.33	0.723	0.00
Signal x Half	3.40	0.066	0.00	0.32	0.572	0.00
Region x Half	0.01	0.923	0.00	0.23	0.629	0.00

Video number x Half	0.05	0.948	0.00	0.08	0.923	0.00
Video condition x Half	0.09	0.762	0.00	0.10	0.756	0.00
Signal x Region x Video number	0.10	0.905	0.00	0.01	0.990	0.00
Signal x Region x Video condition	0.04	0.839	0.00	0.01	0.946	0.00
Signal x Video number x Video condition	0.44	0.644	0.00	0.10	0.908	0.00
Region x Video number x Video condition	0.19	0.829	0.00	0.26	0.773	0.00
Signal x Region x Video Half	0.14	0.710	0.00	0.11	0.744	0.00
Signal x Video number x Half	0.09	0.915	0.00	0.01	0.989	0.00
Region x Video number x Half	0.02	0.981	0.00	0.06	0.938	0.00
Signal x Video condition x Half	<0.001	0.988	0.00	0.56	0.455	0.00
Region x Video condition x Half	0.25	0.616	0.00	0.05	0.828	0.00
Video number x Video condition x Half	1.08	0.341	0.00	0.10	0.906	0.00
Signal x Region x Video number x Video condition	0.06	0.943	0.00	0.28	0.753	0.00
Signal x Region x Video number x Half	0.01	0.994	0.00	0.07	0.932	0.00

Signal x Region x Video condition x Half	0.02	0.903	0.00	0.03	0.865	0.00
Signal x Video number x Video condition x Half	0.004	0.996	0.00	0.26	0.774	0.00
Region x Video number x Video condition x Half	0.24	0.786	0.00	0.10	0.903	0.00
Signal x Region x Video number x Video condition x Half	0.06	0.940	0.00	0.05	0.947	0.00

Table 3.9: Mean, standard deviation and N for average absolute power over video halves for test and retest sessions

Signal	*Region*	*Video number*	*Video condition*	*Video half*	TEST			RETEST		
					M	*SD*	*N*	*M*	*SD*	*N*
Alpha	Frontal	First	Social	First	1.75	0.39	37	1.67	0.44	37
				Second	1.80	0.44	37	1.69	0.50	37
			Non-social	First	1.78	0.43	38	1.65	0.41	37
				Second	1.77	0.39	38	1.67	0.42	37
		Second	Social	First	1.68	0.42	37	1.63	0.37	37
				Second	1.74	0.41	37	1.63	0.41	37
			Non-social	First	1.63	0.33	38	1.65	0.43	36
				Second	1.62	0.32	38	1.63	0.41	36
		Third	Social	First	1.69	0.42	38	1.58	0.50	34
				Second	1.70	0.42	38	1.58	0.45	34

			Non-social	First	1.63	0.41	36	1.62	0.46	34
				Second	1.68	0.40	36	1.60	0.47	34
	Posterior	First	Social	First	1.95	0.46	37	1.99	0.36	37
				Second	2.00	0.48	37	2.06	0.35	37
			Non-social	First	1.98	0.42	38	1.95	0.31	37
				Second	1.94	0.41	38	1.96	0.31	37
		Second	Social	First	1.97	0.56	37	1.96	0.36	37
				Second	1.97	0.50	37	1.99	0.34	37
			Non-social	First	1.85	0.47	38	1.95	0.32	36
				Second	1.88	0.42	38	1.98	0.32	36
		Third	Social	First	2.00	0.42	38	1.95	0.38	34
				Second	1.98	0.45	38	1.97	0.33	34
			Non-social	First	1.88	0.39	36	1.89	0.35	34
				Second	1.93	0.42	36	1.92	0.35	34
Theta	Frontal	First	Social	First	2.42	0.29	37	2.53	0.45	37
				Second	2.54	0.32	37	2.52	0.38	37
			Non-social	First	2.42	0.30	38	2.48	0.36	36
				Second	2.48	0.31	38	2.57	0.35	37
		Second	Social	First	2.47	0.35	37	2.56	0.41	37
				Second	2.58	0.39	37	2.58	0.39	37
			Non-social	First	2.45	0.36	38	2.49	0.33	36

			Second	2.51	0.36	38	2.55	0.42	36
	Third	Social	First	2.51	0.34	38	2.54	0.45	34
			Second	2.57	0.31	38	2.55	0.39	34
		Non-social	First	2.49	0.35	36	2.51	0.35	34
			Second	2.56	0.37	36	2.54	0.38	34
Posterior	First	Social	First	2.51	0.45	37	2.74	0.37	37
			Second	2.67	0.45	37	2.76	0.35	37
		Non-social	First	2.50	0.42	38	2.58	0.34	37
			Second	2.56	0.38	38	2.65	0.40	37
	Second	Social	First	2.60	0.45	38	2.76	0.37	37
			Second	2.70	0.44	38	2.75	0.39	37
		Non-social	First	2.60	0.45	37	2.66	0.32	36
			Second	2.70	0.45	37	2.71	0.42	36
	Third	Social	First	2.71	0.39	38	2.70	0.36	34
			Second	2.71	0.39	38	2.78	0.33	34
		Non-social	First	2.54	0.38	36	2.69	0.35	34
			Second	2.68	0.41	36	2.71	0.34	34

3.3.1.3 Power change: growth curve analysis

To assess power change over the course of video viewing, a linear mixed model approach was used with segment number as a fixed effect and ID as a random effect allowing for random intercept and slope. Investigations revealed there were very little differences between participants' slopes, and this was not a good fit to the data, therefore the model was changed to allow random intercept only. Video number (first, second, third) and video condition (social versus non-social) were additionally included as fixed effects, so that intercepts (i.e. average power at start of video) could be compared across conditions. As average power analyses had revealed differences between theta and alpha power, and power in the frontal and posterior regions (see section 3.3.1), individual models were fitted for each of these.

To check whether the location of the re-reference impacted power, models were rerun including a re-reference variable as a random effect. This was a binary variable indicating whether each trial was referenced to Cz or the average of C3 and C4. For all models, reference location was found to explain only a very small amount of variance and its inclusion in the models did not impact the pattern of results for nearly all models (see appendix B3a). The exceptions were models for frontal and posterior theta test, where the effect of video number became more significant when reference location was included in the model, though the amount of variance explained by this was very small in these models. Given that the location of the reference seemed to have minimal impact on EEG power, further models did not include this variable, as we chose to keep models as simple as possible. The final model (using *lmer* in R (Douglas Bates et al., 2015)) for each analysis was:

Formula: power ~ Segment number + video number + video condition + (1 | ID), data, REML = TRUE)

Intercept values were extracted for each participant, for use in following analyses. Any conditions where a participant provided fewer than ten usable segments of EEG were excluded and assumptions were checked. Analyses were also rerun including a minimum of 30 usable segments of EEG per conditions (see appendix B3a) and intercepts were extracted for reliability checks.

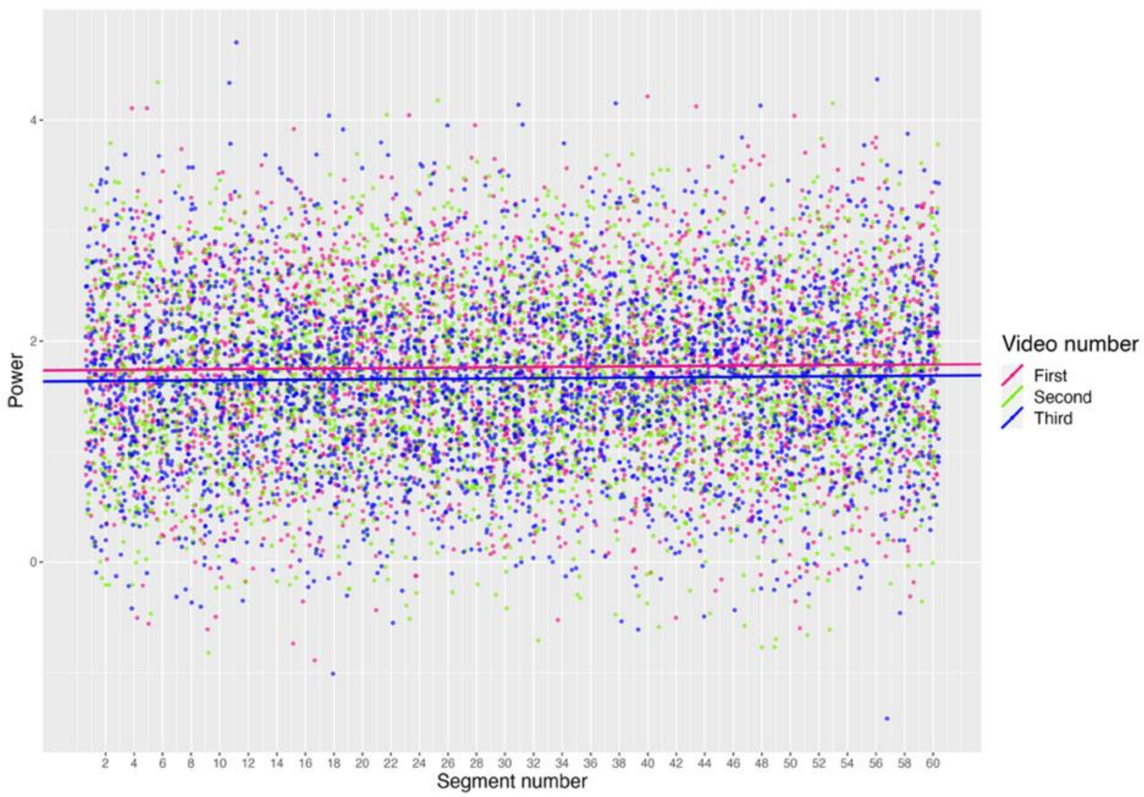
3.3.1.3.1 Test session

Separate models were fitted for frontal and posterior, alpha and theta power. For all models, the intercept corresponds to when segment number is zero, video number is first, and the video condition is non-social.

3.3.1.3.1.1 *Frontal alpha*

45 participants were included with a total of 10144 observations in the model for frontal alpha. The intercept was significantly different from zero *(β* = 1.74*, SE* = 0.06*, t*(53.27) = 30.48, *p* < .001), whilst an ANOVA deviance test conducted on the model found significant effects of segment number and video number but not condition (Table 3.10 and Table 3.12). Pairwise comparisons showed that power was higher in the first versus each of the second and third videos, with no significant differences found between the second and third videos (Table 3.11; Figure 3.8). The estimate of segment number was very small, indicating a very small increase in power over the course of video viewing (Table 3.12 and Table 3.13). Random effects statistics are reported in Table 3.14.

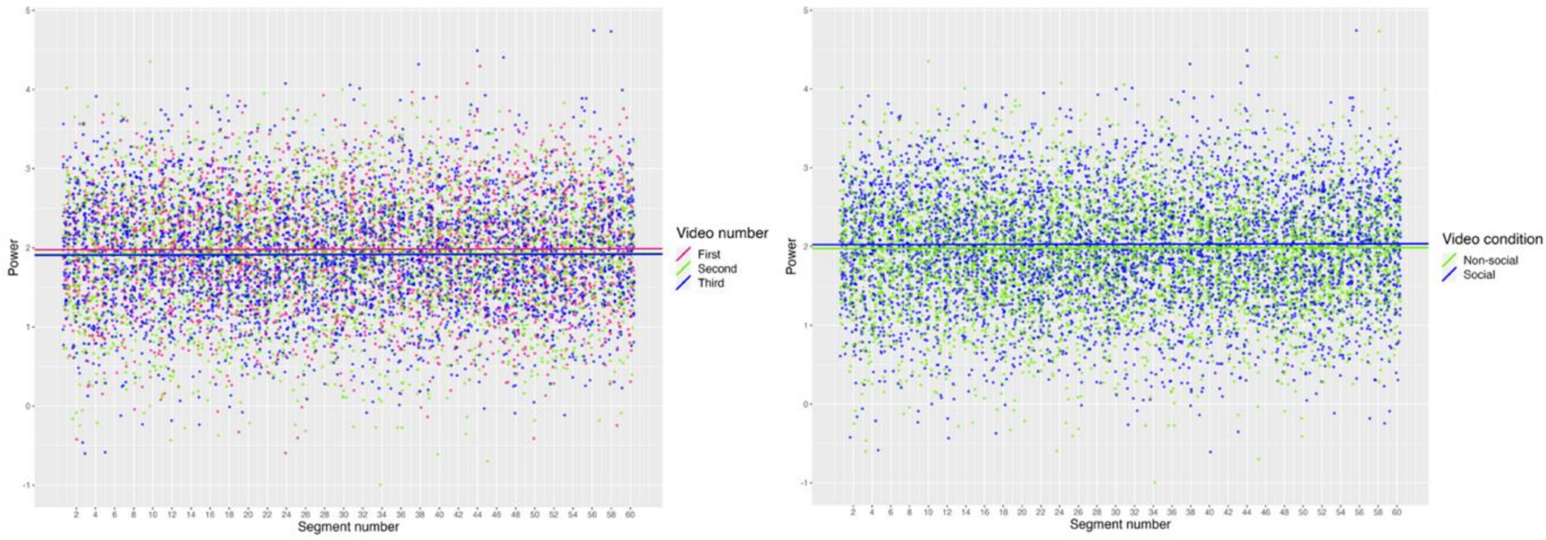
Figure 3.8: Absolute frontal alpha power per segment number with lines for first, second and third estimated from linear mixed model for test session. *Please note the green line for second video is not visible due to nearly identical estimates for the second and third video (see table 3.11)*



3.3.1.3.1.2 Posterior alpha

45 participants were included with a total of 10144 observations in the model for posterior alpha. The intercept was significantly different from zero *(β* = 1.98*, SE* = 0.06*, t*(49.29) = 31.24, *p* < .001), whilst an ANOVA deviance test conducted on the model found a significant effect of condition and video number (Table 3.10 and Table 3.12). Means show that power during social videos was higher than in the non-social condition, and pairwise comparisons showed that power was higher in the first versus each of the second and third videos, with no significant differences found between the second and third videos (Table 3.11; Figure 3.9). The effect of segment number was not significant, indicating no changes in power over the course of video viewing (Table 3.12 and Table 3.13). Random effects statistics are reported in Table 3.14.

Figure 3.9: Absolute posterior alpha power per segment number with lines for (a) first, second and third and (b) social and non-social conditions estimated from linear mixed model for test session



3.3.1.3.1.3 Frontal theta

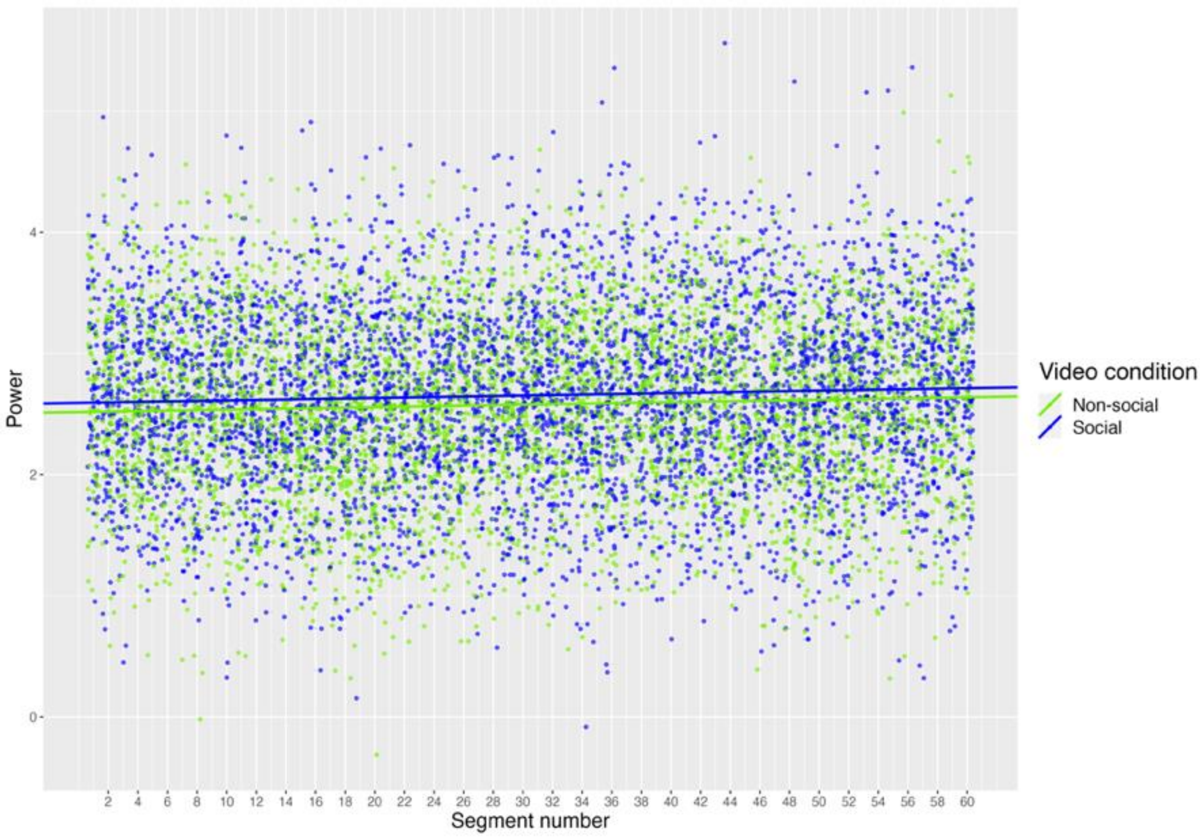
45 participants were included with a total of 10144 observations in the model for frontal theta. The intercept was significantly different from zero *(β* = 2.42*, SE* = 0.05*, t*(49.44) = 43.27, *p* < .001), whilst an ANOVA deviance test conducted on the model found that there were no significant differences between conditions and video

repetitions (Table 3.10 and Table 3.12). The estimate for segment number was small but positive, indicating that theta power increased over the course of video viewing (Table 3.12 and Table 3.13). Random effects statistics are reported in Table 3.14.

3.3.1.3.1.4 Posterior theta

45 participants were included with a total of 10144 observations in the model for posterior theta. The intercept was significantly different from zero *(β* = 2.52*, SE* = 0.06*, t*(50.36) = 41.11, *p* < .001), whilst an ANOVA deviance test conducted on the model found significant effects of segment number and condition but not video number (Table 3.10 and Table 3.12). Power was significantly higher in the social versus non-social condition (Figure 3.10), and the estimate for segment number was small but positive, indicating that theta power increased over the course of video viewing (Table 3.12 and Table 3.13). Random effects statistics are reported in Table 3.14.

Figure 3.10: Absolute frontal theta power per segment number with lines for social and non-social conditions estimated from linear mixed model for test session

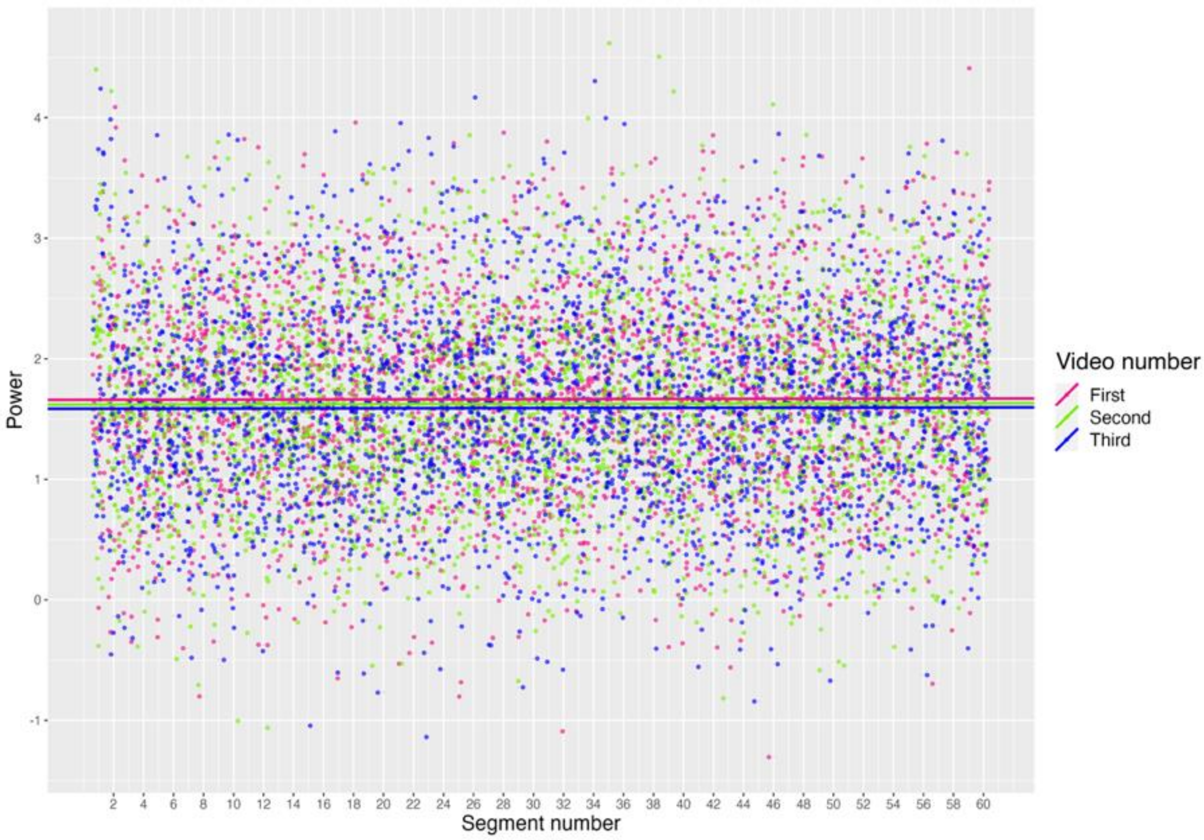


3.3.1.3.2 Retest session

3.3.1.3.2.1 Frontal alpha

The intercept was significantly different from zero *(β* = 1.66*, SE* = 0.06*, t*(44.96) = 25.90, *p* < .0001), whilst an ANOVA deviance test conducted on the model found a significant effect of video number but not condition or segment number (Table 3.10 and Table 3.12). Pairwise comparisons showed that power was higher in the first versus third video but there was no difference between first and second, or second and third videos; (Table 3.11; Figure 3.11). The effect of segment number was not significant, indicating no changes in power over the course of video viewing (Table 3.12 and Table 3.13). Random effects statistics are reported in Table 3.14.

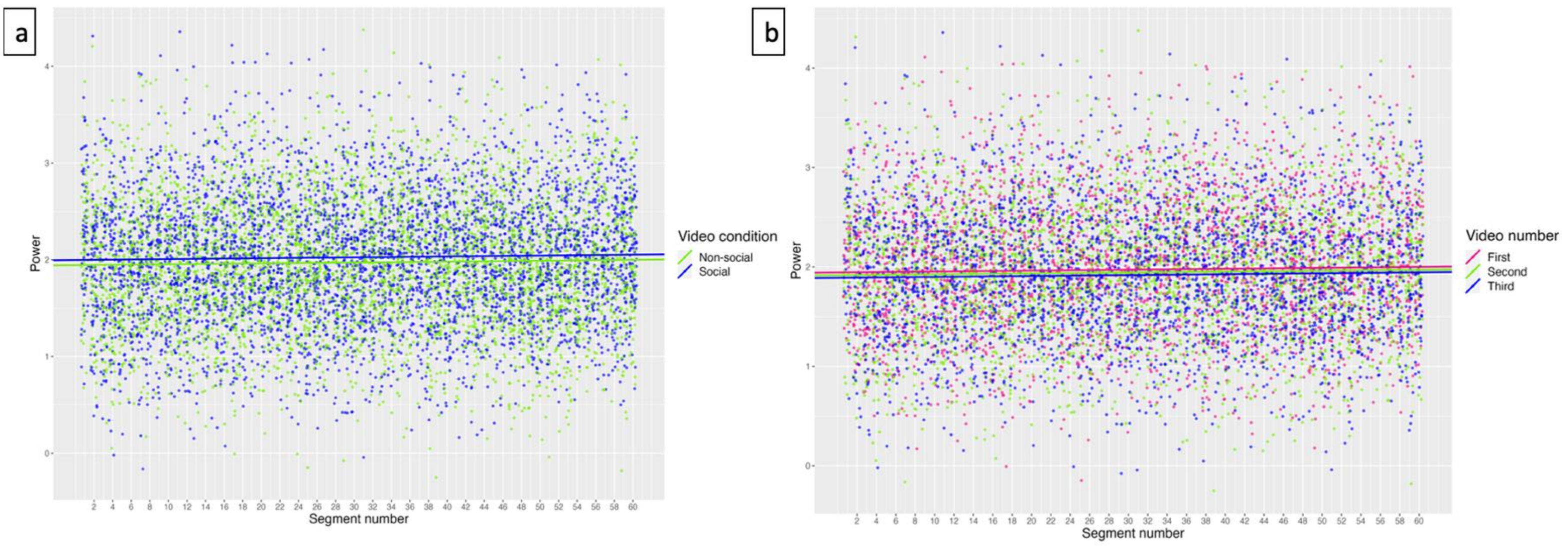
Figure 3.11: Absolute frontal alpha power per segment number with lines for first, second and third estimated from linear mixed model for retest session



3.3.1.3.2.2 Posterior alpha

The intercept was significantly different from zero *(β* = 1.94*, SE* = 0.05*, t*(45.72) = 38.19, *p* < .0001), whilst an ANOVA deviance test conducted on the model found significant effects of segment number, condition, and video number. Means showed that power was lower in the non-social versus social condition (Table 3.10 and Table 3.12), whilst pairwise comparisons showed that power was higher in the first versus third video but there no difference between first and second, or second and third videos (Table 3.11; Figure 3.12). The estimate for segment number was very small but positive, indicating that theta power increased by a small degree over the course of video viewing (Table 3.12 and Table 3.13). Random effects statistics are reported in Table 3.14.

Figure 3.12: Absolute posterior alpha power per segment number with lines for (a) social and non-social conditions and (b) first, second and third estimated from linear mixed model for retest session

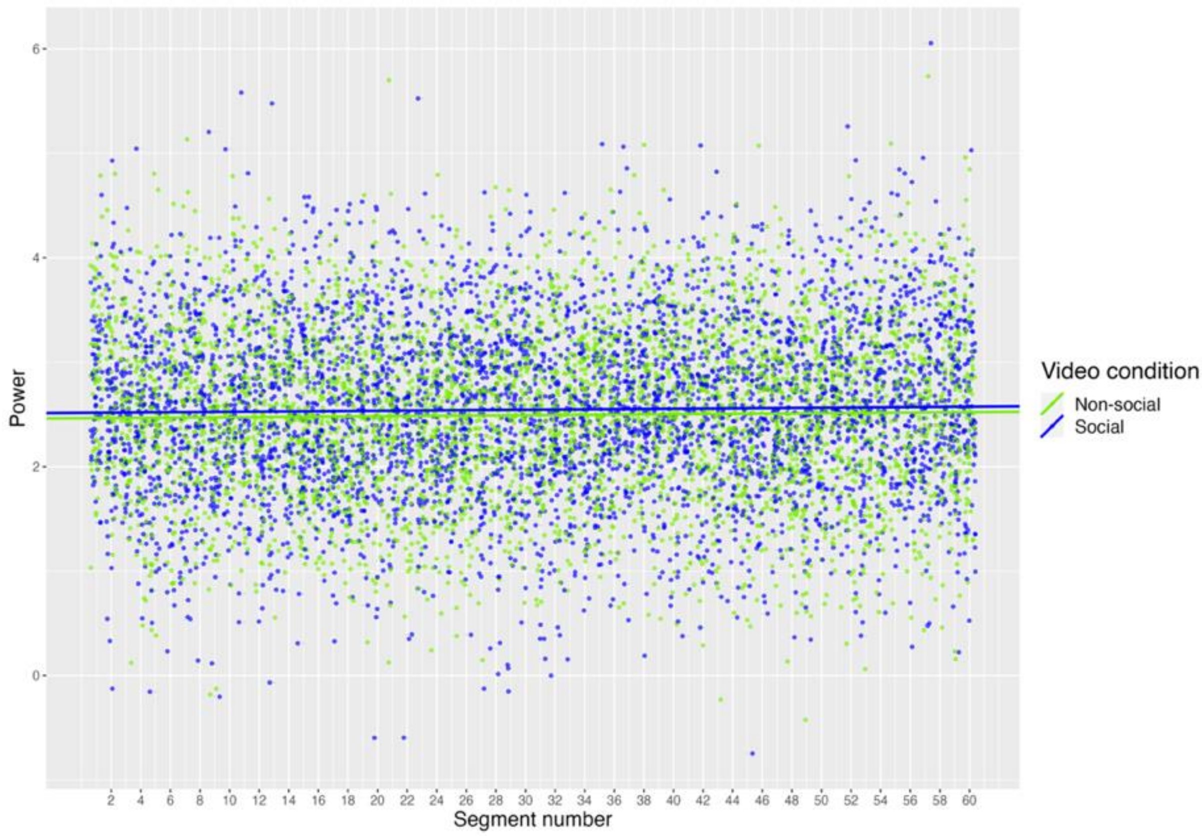


3.3.1.3.2.3 Frontal theta

39 participants were included with a total of 9351 observations in the model for frontal theta. The intercept was significantly different from zero *(β* = 2.46*, SE* = 0.06*, t*(48.43) = 43.27, *p* < .0001), whilst an ANOVA deviance test conducted on the model found significant effects for segment number and condition but not video number (Table 3.10 and Table 3.12). Means showed that power was lower in the non-social

versus social condition (Figure 3.13). The estimate for segment number was very small but positive, indicating that theta power increased by a small degree over the course of video viewing (Table 3.12 and Table 3.13). Random effects statistics are reported in Table 3.14.

Figure 3.13: Absolute frontal theta power per segment number with lines for social and non-social estimated from linear mixed model for retest session



3.3.1.3.2.4 Posterior theta

The intercept was significantly different from zero *(β* = 2.61*, SE* = 0.05, *t*(46.11) = 48.68, *p* < .0001), whilst an ANOVA deviance test conducted on the model found significant effects of condition and segment number, but not video repetitions. Means showed that power was higher in the social versus non-social condition (Table 3.10 and Table 3.12; Figure 3.14). The estimate for segment number was small but positive, indicating that theta power increased by a small degree over the course of video viewing (Table 3.12 and Table 3.13). Random effects statistics are reported in Table 3.14.

Figure 3.14: Absolute posterior theta power per segment number with lines for social and non-social conditions estimated from linear mixed model for retest session

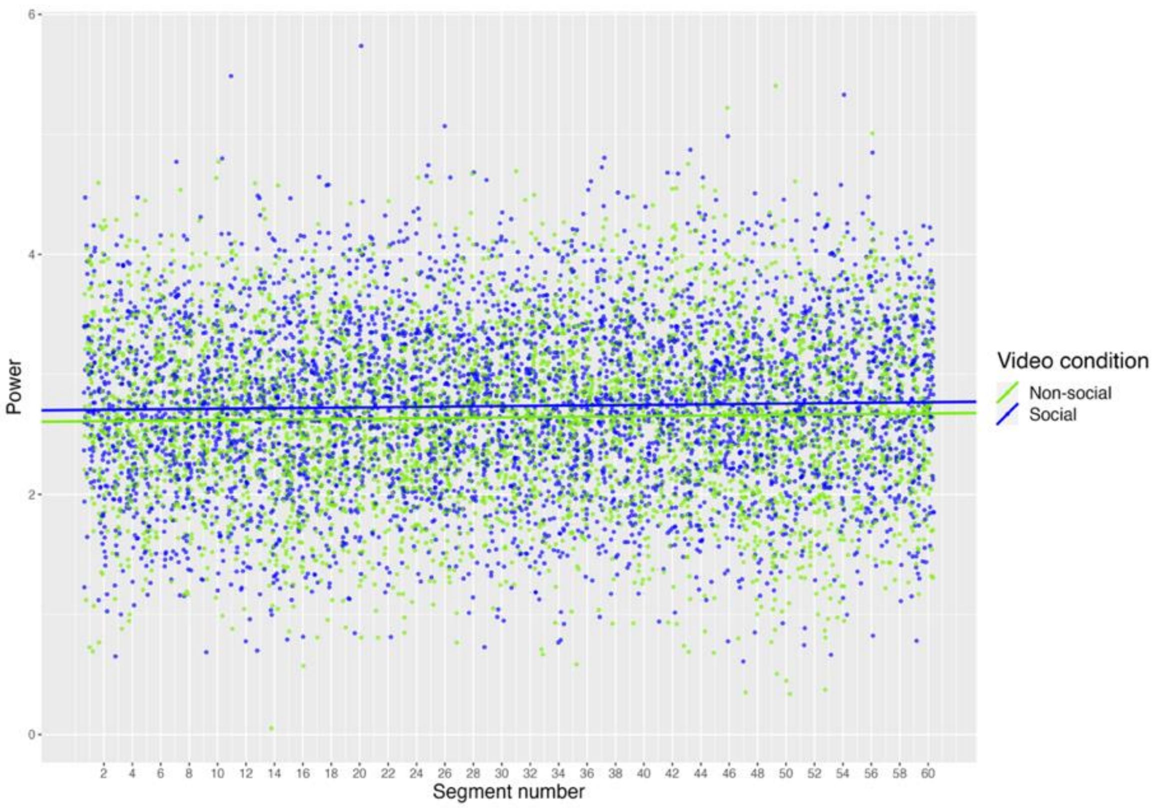


Table 3.10: Mean, standard error and 95% confidence intervals for first, second and third videos, and social and non-social conditions from average absolute power per segment in the test and retest session

Signal		TEST				RETEST			
		M	*SE*	95% CIs		*M*	*SE*	95% CIs	
		Video number							
	First	1.77	0.06	1.66	1.88	1.67	0.06	1.54	1.80
	Second	1.67	0.06	1.56	1.78	1.63	0.06	1.51	1.76
Frontal alpha	Third	1.67	0.06	1.56	1.78	1.60	0.06	1.47	1.72
		Video condition							
	Social	0.29	0.01	0.27	0.31	0.27	0.01	0.25	0.29
	Non-social	0.28	0.01	0.26	0.30	0.27	0.01	0.25	0.29
		Video number							
	First	2.01	0.06	1.88	2.13	2.00	0.05	1.90	2.10
	Second	1.95	0.06	1.82	2.07	1.97	0.05	1.87	2.07
Posterior alpha	Third	1.94	0.06	1.81	2.06	1.94	0.05	1.85	2.04
		Video condition							
	Social	1.99	0.06	1.86	2.11	2.00	0.05	1.90	2.10
	Non-social	1.94	0.06	1.81	2.06	1.95	0.05	1.85	2.04
		Video number							
	First	2.50	0.05	2.40	2.59	2.52	0.06	2.41	2.63
	Second	2.51	0.05	2.42	2.60	2.53	0.06	2.42	2.64
Frontal theta	Third	2.54	0.05	2.45	2.63	2.52	0.06	2.41	2.63
		Video condition							
	Social	2.53	0.05	2.44	2.62	2.55	0.06	2.44	2.66
	Non-social	2.50	0.05	2.41	2.59	2.50	0.06	2.39	2.61

				Video number					
	2.62	0.06	2.50	2.74	2.69	0.05	2.58	2.79	2.62
	2.65	0.06	2.53	2.77	2.72	0.05	2.61	2.82	2.65
Posterior theta	2.65	0.06	2.53	2.77	2.72	0.05	2.61	2.82	2.65
				Video condition					
	2.68	0.06	2.56	2.79	2.75	0.05	2.65	2.86	2.68
	2.60	0.06	2.48	2.72	2.66	0.05	2.56	2.76	2.60

Table 3.11: Estimate, standard error, t-value, and p-value for pairwise comparisons of average absolute power per segment number between video numbers for frontal and posterior alpha in test and retest conditions. Significant effects are indicated with a *

			TEST				RETEST			
			β	*SE*	*t*	*p*	β	*SE*	*t*	*p*
Frontal alpha	First	Second	0.10	0.02	2.87	< .0001*	0.04	0.02	2.24	.08
	First	Third	0.10	0.02	5.81	< .0001*	0.07	0.02	4.11	.0001*
	Second	Third	-0.002	0.02	-0.09	1.00	0.03	0.02	1.81	.17
Posterior alpha	First	Second	0.06	0.02	4.06	.0001*	0.03	0.02	1.89	.18
	First	Third	0.07	0.02	4.78	< .0001*	0.05	0.02	3.63	.001*
	Second	Third	0.01	0.01	0.71	1.00	0.03	0.02	1.77	.23

Table 3.12: Chi-squared, p-value and partial eta squared for segment number, video number and condition effects from an ANOVA of linear mixed models of average absolute power per segment number, video number and condition for frontal and posterior alpha and theta for the test and retest sessions. Significant effects are indicated with a *

		TEST			RETEST		
		$\chi 2$	*p*	η^2 (partial)	$\chi 2$	*p*	η^2 (partial)
Frontal alpha	Segment number	4.63	.03*	<.001	0.18	.68	<.001
	Video number	44.69	< .0001*	0.004	16.94	< .001*	0.002
	Condition	2.28	.13	<.001	0.55	.46	<.001
Posterior alpha	Segment number	0.33	.57	<.001	6.86	.009*	<.001
	Video number	26.16	< .001*	0.003	13.19	.001*	0.001
	Condition	16.18	< .001*	0.002	19.54	< .0001*	0.002
Frontal theta	Segment number	26.45	< .0001*	0.003	4.73	.03*	<.001
	Video number	5.96	.51	<.001	0.83	.66	<.001
	Condition	3.24	.07	<.001	11.57	.0007*	0.001
Posterior theta	Segment number	32.13	< .001*	0.003	8.53	.004*	<.001
	Video number	5.84	.05	<.001	4.06	.13	<.001
	Condition	36.18	< .001*	0.004	53.49	< .001*	0.006

Table 3.13: Estimate, standard error, degrees of freedom and 95% confidence intervals for each of the fixed effects in linear mixed models performed on average absolute power per segment for each of frontal and posterior alpha and theta in the test and retest sessions

Fixed Effects										
	TEST					RETEST				
	β	*SE*	*DF*	CIs 95%		*β*	*SE*	*DF*	CIs 95%	
Frontal alpha										
Intercept	1.74	0.06	53.27	1.67	1.86	1.66	0.06	44.96	1.57	1.75
Segment Number	<.001	<.001	10095.79	<.001	0.002	<.001	<.001	9308.71	<.001	0.002
Video number: second	-0.10	0.02	10121.00	-0.13	-0.05	-0.04	0.02	9317.34	-0.06	-0.02
Video number: third	-0.10	0.02	10126.47	-0.14	-0.07	-0.07	0.02	9315.71	-0.11	-0.06
Video condition: social	0.02	0.01	10122.63	-0.002	0.05	0.01	0.01	9314.17	-0.01	0.04
Posterior alpha										
Intercept	1.98	0.06	49.29	1.90	2.11	1.94	0.05	45.72	1.87	2.01
Segment Number	<.001	<.001	10095.62	-0.0001	0.001	0.001	<.001	9308.84	<-.001	0.002
Video number: second	-0.06	0.02	1011.34	-0.08	-0.02	-0.03	0.02	9318.18	-0.05	-0.01
Video number: third	-0.07	0.02	10115.53	-0.10	-0.05	-0.05	0.02	9316.44	-0.08	-0.05
Video condition: social	0.05	0.01	10112.58	0.03	0.07	0.05	0.01	9314.77	0.04	0.08
Frontal theta										
Intercept	2.42	0.05	59.01	2.37	2.52	2.46	0.06	48.53	2.38	2.53
Segment Number	0.002	<.001	10096.03	0.002	0.003	0.001	<.001	9308.96	.0001	0.003
Video number: second	0.01	0.02	10130.70	-0.01	0.06	0.02	0.02	9321.26	-0.01	0.04

Video number: third	0.04	0.02	10135.67	0.01	0.07	<.001	0.02	9319.05	-0.04	0.01
Video condition: social	0.03	0.02	10132.25	<0.001	0.05	0.05	0.02	9316.84	0.03	0.08
Posterior theta										
Intercept	2.52	0.06	50.36	2.45	2.65	2.61	0.05	46.11	2.53	2.68
Segment Number	0.002	<.001	10095.64	0.002	0.003	0.001	<.001	2.92	<.001	0.002
Video number: second	0.03	0.02	10114.28	0.01	0.08	0.03	0.02	1.76	0.01	0.05
Video number: third	0.03	0.02	10118.99	0.003	0.06	0.03	0.02	1.73	-0.003	0.04
Video condition: social	0.08	0.01	10115.68	0.05	0.10	0.09	0.01	7.31	0.08	0.12

Table 3.14: Estimate and standard error for each of the random effects in linear mixed models performed on average absolute power per segment for each of frontal and posterior alpha and theta in the test and retest sessions

Random Effects					
		TEST		RETEST	
		β	*SE*	β	*SE*
Frontal alpha	ID: Intercept	0.13	0.36	0.15	0.38
	Residual	0.47	0.69	0.48	0.69
Posterior alpha	ID: Intercept	0.17	0.41	0.09	0.30
	Residual	0.34	0.58	0.33	0.57
Frontal theta	ID: Intercept	0.09	0.30	0.11	0.33
	Residual	0.53	0.73	0.54	0.74
Posterior theta	ID: Intercept	0.15	0.39	0.10	0.32
	Residual	0.38	0.62	0.38	0.62

3.3.1.4 Variability

Table 3.15 shows descriptive data for variability measures per signal (alpha and theta) and region (frontal and posterior). Residual standard errors were calculated in relation to regression lines estimated by the linear mixed models described in section 3.3.1.3 and are closely related to standard error of estimates, which were calculated using the 'se.ranef' function in R (Douglas Bates et al., 2015). Standard deviations were calculated in relation to mean EEG power when averaged over the whole video.

Table 3.15: Mean, standard deviation and number of participants for residual standard error, standard error of estimates and standard deviation of absolute alpha and theta power in frontal and posterior regions for test and retest sessions

Variability statistic			TEST			RETEST		
	Signal	*Region*	*M*	*SD*	*N*	*M*	*SD*	*N*
Residual standard error	Alpha	Frontal	0.47	0.08	45	0.49	0.11	39
		Posterior	0.34	0.06		0.33	0.07	
	Theta	Frontal	0.53	0.08		0.55	0.08	
		Posterior	0.38	0.06		0.38	0.06	
Standard error of estimates	Alpha	Frontal	0.05	0.01	45	0.05	0.01	39
		Posterior	0.04	0.01		0.04	0.01	
	Theta	Frontal	0.05	0.01		0.05	0.01	
		Posterior	0.04	0.01		0.04	0.01	
Standard deviation	Alpha	Social	0.68	0.06	45	0.70	0.07	39
		Non-social	0.58	0.05		0.57	0.06	
	Theta	Social	0.73	0.06		0.74	0.05	
		Non-social	0.62	0.05		0.62	0.04	

3.3.2 Relative power

3.3.2.1 Average power

To compare how average power differed across conditions within each of the test and retest sessions, repeated-measures linear models were fitted using the *lm* function in R (Core Team, 2023). Average power over the whole 60 seconds of video viewing was the dependent variable and independent variables of EEG signal (alpha versus theta), region (frontal or posterior), video number (first, second or third) and video condition (social versus non-social). Videos which had fewer than ten usable segments of EEG data were excluded from analyses and assumptions were checked. Analogous analyses conducted using a minimum of 30 usable EEG segments per condition found the same pattern of findings (see appendix B1b). Identical models were fitted for data from the test and retest session; where significant interactions were found, follow-up pairwise contrasts were conducted using the *emmeans* function in R (Lenth, 2023) with a Bonferroni correction for p-values.

3.3.2.1.1 Test session

Table 3.20 show statistics from an ANOVA run on the linear model fit with relative power data from the test session using *anova* in R (Fox & Weisberg, 2019); Table 3.21 shows related descriptive statistics. Significant results were found for the main effects of signal, region, and video condition, and for the interactions between signal and region, and signal and video number.

Means indicated that average theta power (*M* = 0.41, *SE* = 0.003, CIs [0.40, 41])) was higher than average alpha power (*M* = 0.28, *SE* = 0.003, CIs [0.28, 0.29]), average power was higher in the frontal (*M* = 0.36, *SE* = 0.003, CIs [0.35, 0.36]) compared to posterior (*M* = 0.33, *SE* = 0.003, CIs [0.33, 0.34]) region, and in the social (*M* = 0.35, *SE* = 0.003, CIs [0.35, 0.36]) compared to non-social (*M* = 0.34, *SE* = 0.003, CIs [0.33, 0.34]) condition. Pairwise comparisons found no significant differences in power between the three videos (see Table 3.16 and Table 3.17).

Table 3.16: Estimate, t-value, p-value, and standard errors for pairwise comparisons between average relative power in the first, second and third videos in the test session

Effect	Effect	*β*	*t*	*p*	*SE*
First	Second	0.011	2.16	.078	0.01
First	Third	0.012	2.29	.056	0.01
Second	Third	0.001	0.12	.992	0.01

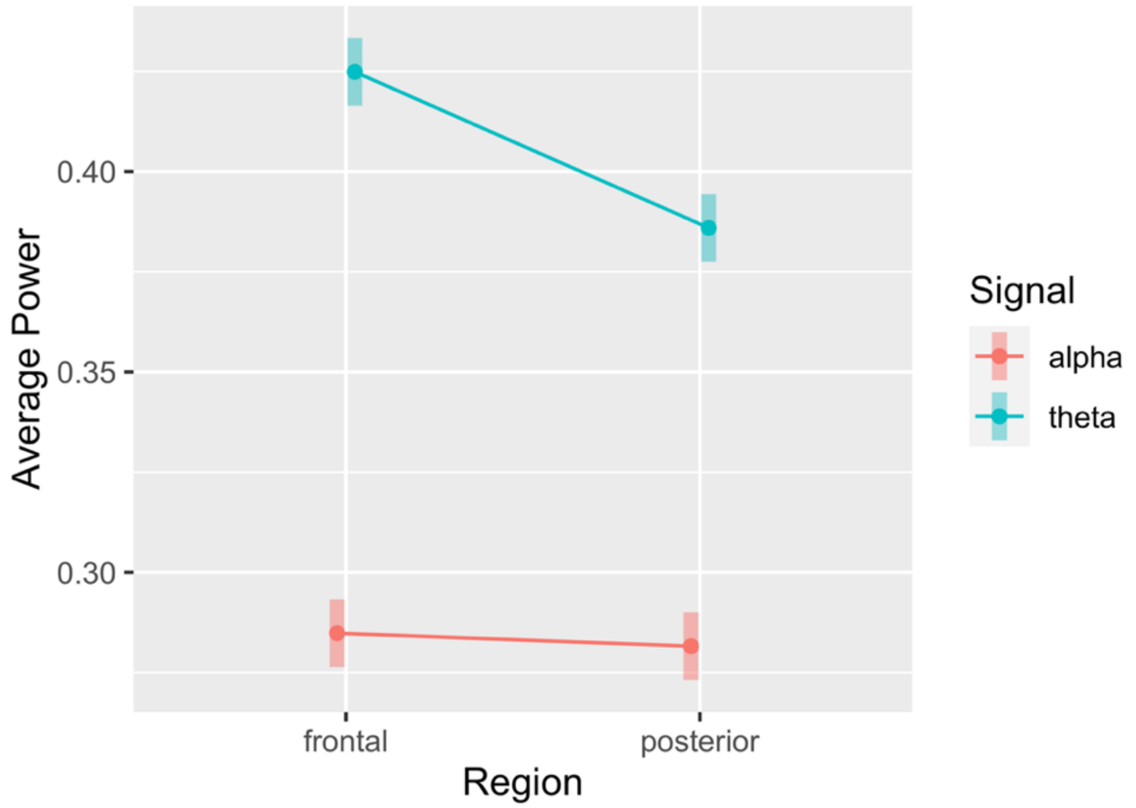
Table 3.17: Mean, standard error, and 95% confidence intervals for average relative power in the first, second and third videos, averaged from first and second half of video viewing in the test session

Video condition	*M*	*SE*	95% CIs	
First	0.35	0.004	0.35	0.36
Second	0.34	0.004	0.33	0.35
Third	0.34	0.004	0.33	0.35

Signal x region

Contrasts indicated that average theta power in the frontal (*M* = 0.43, *SE* = 0.004, [0.42, 0.43]) was significantly higher compared to the posterior region (*M* = 0.39, *SE* = 0.004, [0.38, 0.39]) [β = 0.04, t(900) = 6.41, *SE* = 0.01, *p* < .0001], though no differences were found between average alpha power in the frontal (*M* = 0.29, *SE* = 0.004, [0.28, 0.29]) compared to posterior (mean = 0.28, *SE* = 0.004, [0.27, 0.29]) region, [β = 0.003, t(900) = 0.53, *SE* = 0.01, *p* = .951]. Posterior alpha was significantly lower than posterior theta power, [β = -0.10, t(900) = -17.18, *SE* = 0.01, *p* < .0001], and frontal alpha power was significantly lower than frontal theta power, [β = -0.14, t(900) = - 23.05, *SE* = 0.01, *p* < .0001]; Figure 3.15.

Figure 3.15: Mean and confidence intervals for relative frontal and posterior alpha and theta power over the whole video in the test session



Signal x video number

Contrasts showed that average alpha power was higher in the first compared to the second, [t(824) = 3.36, *SE* = 0.01, *p* = .010], and third, [t(824) = 3.99, *SE* = 0.01, *p* = .001], video, though no significant differences were found between the second and third videos, [t(0.60), *SE* = 0.01, *p* = .991]; Figure 3.16. No differences were found in average theta power across the videos (Table 3.18 and Table 3.19).

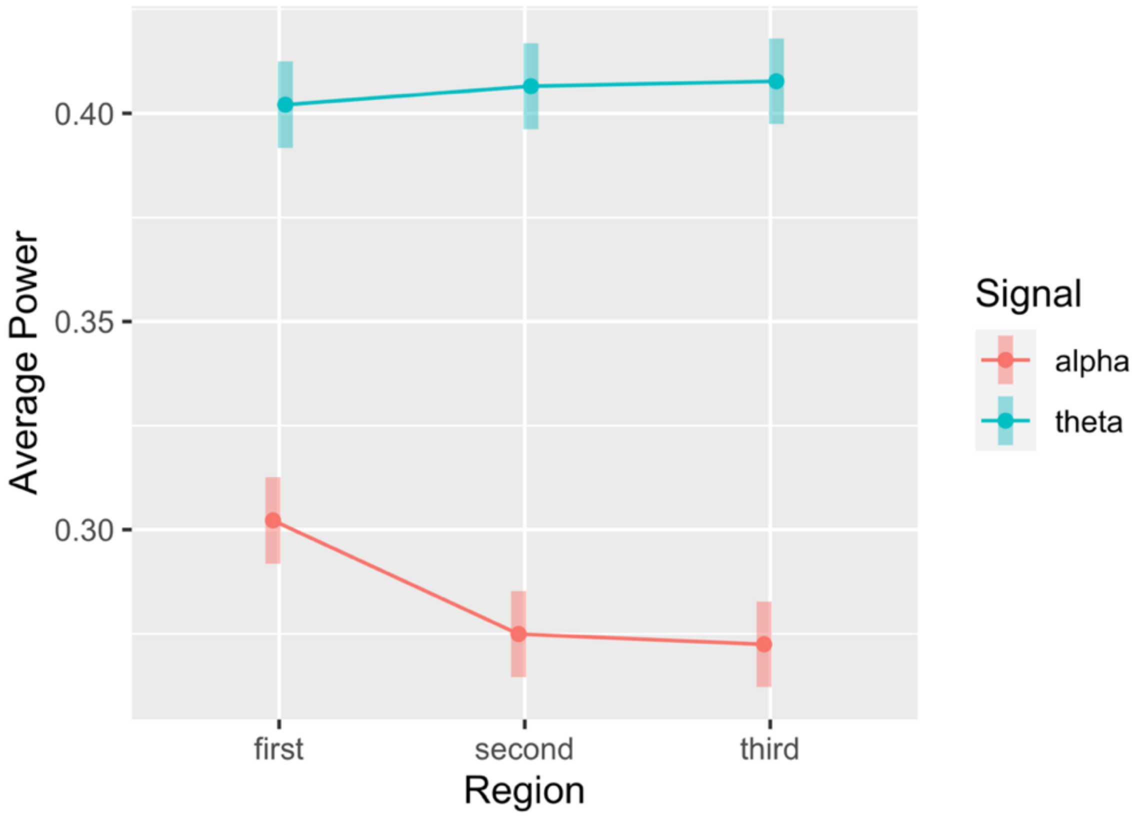
Table 3.18: Estimate, t-value, p-value, and standard errors for pairwise comparisons between average relative alpha and theta power over the whole video in the first, second and third videos in the test session. Significant effects are indicated with a *

Effect	Effect	Estimate	*t*	*p*	*SE*
First, alpha	Second, alpha	0.028	3.65	.004*	0.01
Second, alpha	Third, alpha	0.030	4.00	.001*	0.01
Second, alpha	Third, alpha	0.003	0.33	1.000	0.01
First, theta	Second, theta	-0.004	-0.60	.991	0.01
First, theta	Third, theta	-0.006	-0.76	.975	0.01
Second, theta	Third, theta	-0.001	-0.16	1.00	0.01
First, alpha	First, theta	-0.100	-13.33	< .0001*	0.01
Second, alpha	Second, theta	-0.13	-17.68	< .0001*	0.01
Third, alpha	Third, theta	-0.14	-18.29	< .0001*	0.01

Table 3.19: Mean, standard error, and 95% confidence intervals for relative alpha and theta power in the first, second and third videos in the test session

Signal	*Video condition*	*M*	*SE*	95% CIs	
Alpha	First	0.30	0.01	0.29	0.31
	Second	0.28	0.01	0.27	0.29
	Third	0.27	0.01	0.26	0.28
Theta	First	0.40	0.01	0.39	0.41
	Second	0.41	0.01	0.40	0.42
	Third	0.41	0.01	0.40	0.42

Figure 3.16: Mean and confidence intervals for relative alpha and theta power in the first, second and third videos over the whole video in the test session



3.3.2.1.2 Retest session

Table 3.20 show statistics from an ANOVA run on the linear model fit with relative power data from the retest session using *anova* in R (Fox & Weisberg, 2019); Table 3.21 shows related descriptive statistics. Similarly to in the test session, significant main effects were found for signal, region and video condition and a significant interaction was found between signal and region. In contrast to results from the test session, a significant main effect of video number and a significant interaction between signal and video number were not found in the retest session (Table 3.20 and Table 3.21).

Means indicated that average theta power (*M* = 0.42, *SE* = 0.003, CIs [0.42, 0.43]) was higher than average alpha power (*M* = 0.27, *SE* = 0.003, CIs [0.26, 0.28]), average power was higher in the frontal (*M* = 0.35, *SE* = 0.003, CIs [0.35, 0.36]) compared to posterior (*M* = 0.34, *SE* 0.003, CIs [0.33, 0.34]) region, and in the social

(*M* = 0.35, *SE* = 0.003, CIs [0.35, 0.36]) compared to non-social (*M* = 0.34, *SE* = 0.003, CIs [0.33, 0.35]) condition.

Signal x region

Contrasts indicated that there was no difference between average alpha power in the frontal (*M* = 0.27, *SE* = 0.004, CIs [0.26, 0.28]) compared to the posterior region (*M* = 0.27, *SE* = 0.004, CIs [0.26, 0.28]) [β = <0.001, t(848) = -0.06, *SE* = 0.01, p < 1.00], but theta power was higher in the frontal (*M* = 0.44, *SE* = 0.004, [0.43, 0.45]) compared to posterior (*M* = 0.41, *SE* = 0.004, CIs [0.40, 0.41]) region [β = 0.032, *t*(848) = 5.17, *SE* = 0.01, p < .001]. Posterior alpha power was significantly lower than posterior theta power, [β = -0.136, *t*(848) = -22.08, *SE* = 0.01, p < .0001], and frontal alpha power was significantly lower than frontal theta power, [β = -0.168, *t*(848) = - 27.31, *SE* = 0.01, p < .0001]; Figure 3.17.

Figure 3.17: Mean and confidence intervals for relative frontal and posterior alpha and theta power over the whole video in the retest session

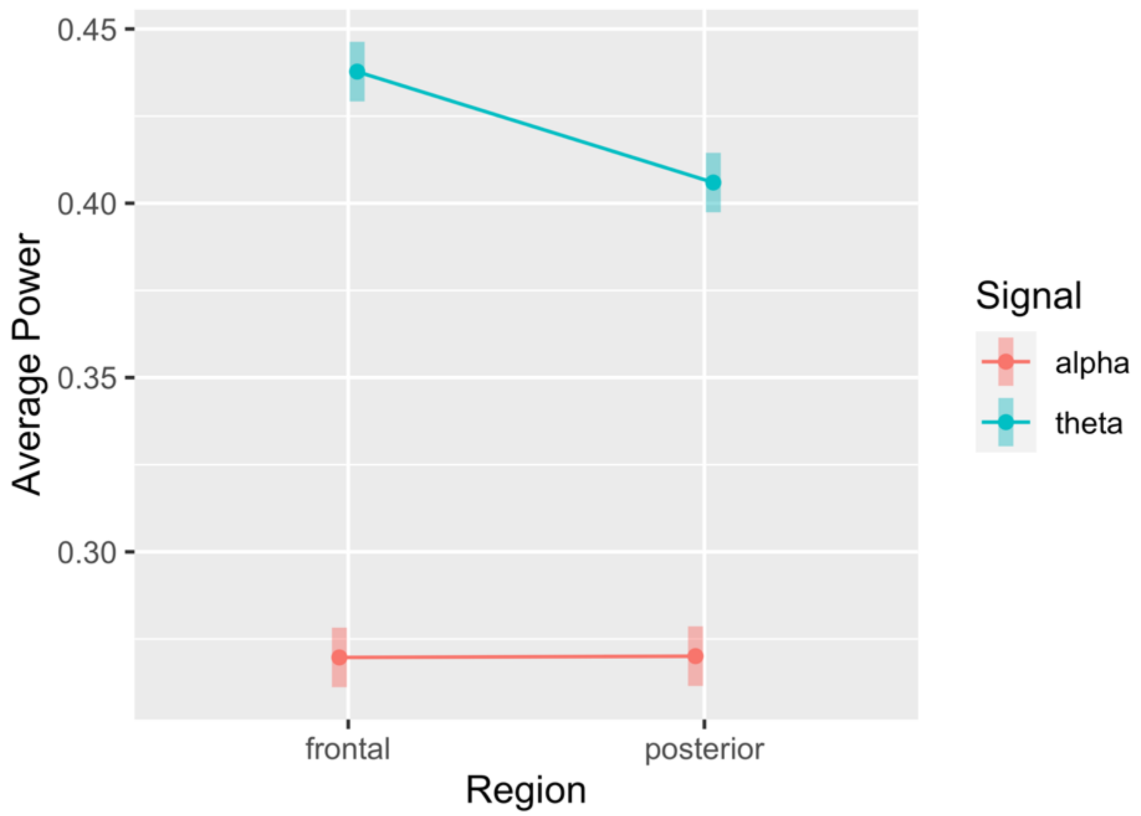


Table 3.20: F-value, p-value and partial eta squared effect size for ANOVAs on average relative power over the whole video for test and retest sessions. Significant effects are indicated with a *

	TEST			RETEST		
	F	*p*	η^2 (partial)	*F*	*p*	η^2 (partial)
Signal	811.21	< .001*	0.47	1216.29	< .001*	0.59
Region	24.01	< .001*	0.03	13.23	.0003*	0.02
Video number	3.37	.035*	0.00	2.60	.075	0.00
Video condition	10.50	.001*	0.01	6.33	.012*	0.00
Video condition x Video number	1.01	.364	0.00	0.44	.646	0.00
Signal x Region	17.24	< .001*	0.02	13.82	.0002*	0.02
Signal x Video number	6.88	.001*	0.02	2.29	.102	0.00
Signal x Video condition	< 0.001	.999	0.00	0.61	.437	0.00
Region x Video number	0.16	.849	0.00	0.42	.655	0.00
Region x Video condition	0.91	.340	0.00	< 0.001	.998	0.00
Region x Video condition x Video number	0.28	.756	0.00	0.16	.852	0.00
Signal x Video condition x Video number	0.83	.439	0.00	0.01	.995	0.00
Signal x Region x Video number	0.17	.848	0.00	0.10	.907	0.00
Signal x Region x Video condition	0.02	.896	0.00	0.71	.400	0.00
Signal x Region x Video number x Video condition	0.07	0.934	0.00	0.45	.641	0.00

Table 3.21: Mean, standard deviation and N for average relative power over the whole video for test and retest sessions

Signal	*Region*	*Video number*	*Video condition*	TEST			RETEST		
				M	*SD*	*N*	*M*	*SD*	*N*
Alpha	Frontal	First	Social	0.31	0.08	37	0.29	0.07	37
			Non-social	0.31	0.08	38	0.28	0.07	38
		Second	Social	0.29	0.07	38	0.27	0.07	36
			Non-social	0.26	0.07	39	0.26	0.07	36
		Third	Social	0.28	0.08	40	0.26	0.07	35
			Non-social	0.27	0.07	38	0.26	0.08	35
	Posterior	First	Social	0.30	0.06	37	0.29	0.06	37
			Non-social	0.30	0.07	39	0.28	0.06	38
		Second	Social	0.29	0.07	38	0.27	0.07	37
			Non-social	0.26	0.07	39	0.26	0.06	36
		Third	Social	0.28	0.07	40	0.27	0.06	35
			Non-social	0.26	0.06	38	0.26	0.06	35
Theta	Frontal	First	Social	0.43	0.06	37	0.45	0.07	37
			Non-social	0.41	0.06	39	0.44	0.06	38
		Second	Social	0.43	0.07	38	0.44	0.08	37
			Non-social	0.42	0.06	39	0.42	0.07	36
		Third	Social	0.43	0.06	40	0.45	0.06	35
			Non-social	0.43	0.06	38	0.43	0.05	35

Posterior	First	Social	0.39	0.05	37	0.42	0.06	38
		Non-social	0.38	0.06	39	0.39	0.05	38
	Second	Social	0.40	0.06	38	0.41	0.07	37
		Non-social	0.38	0.06	39	0.40	0.06	36
	Third	Social	0.40	0.06	40	0.41	0.06	35
		Non-social	0.37	0.05	38	0.41	0.05	35

3.3.2.2 Power change: halves

To compare how average power differed over the course of video viewing, repeated-measures linear models were fitted using the *lm* function in R (Core Team, 2023) using average power in each of the first and second 30 seconds of video viewing as the dependent variable. Independent variables were EEG signal (alpha versus theta), region (frontal or posterior), video number (first, second or third), video condition (social versus non-social) and video half (first or second). Videos were excluded from analyses if either half had fewer than five usable segments of EEG and assumptions were checked. Robustness checks including a minimum of 15 usable EEG segments per half did not find a significantly different pattern of results (see appendix B2b). Identical models were fitted for data from the test and retest session; where significant interactions were found, follow-up pairwise contrasts were conducted using the *emmeans* function in R (Fox & Weisberg, 2019) with a Bonferroni correction for p-values.

3.3.2.2.1 Test session

A significant main effect of halves was not found, but the interaction between signal and video half was significant (Table 3.24 and Table 3.25).

Signal x half

Average theta power was significantly higher in the second half compared to the first half of video viewing, though no differences were found in alpha power (Table 3.22 and Table 3.23; Figure 3.18).

Figure 3.18: Mean and confidence intervals for relative alpha and theta power in the first and second half of video viewing for the test session

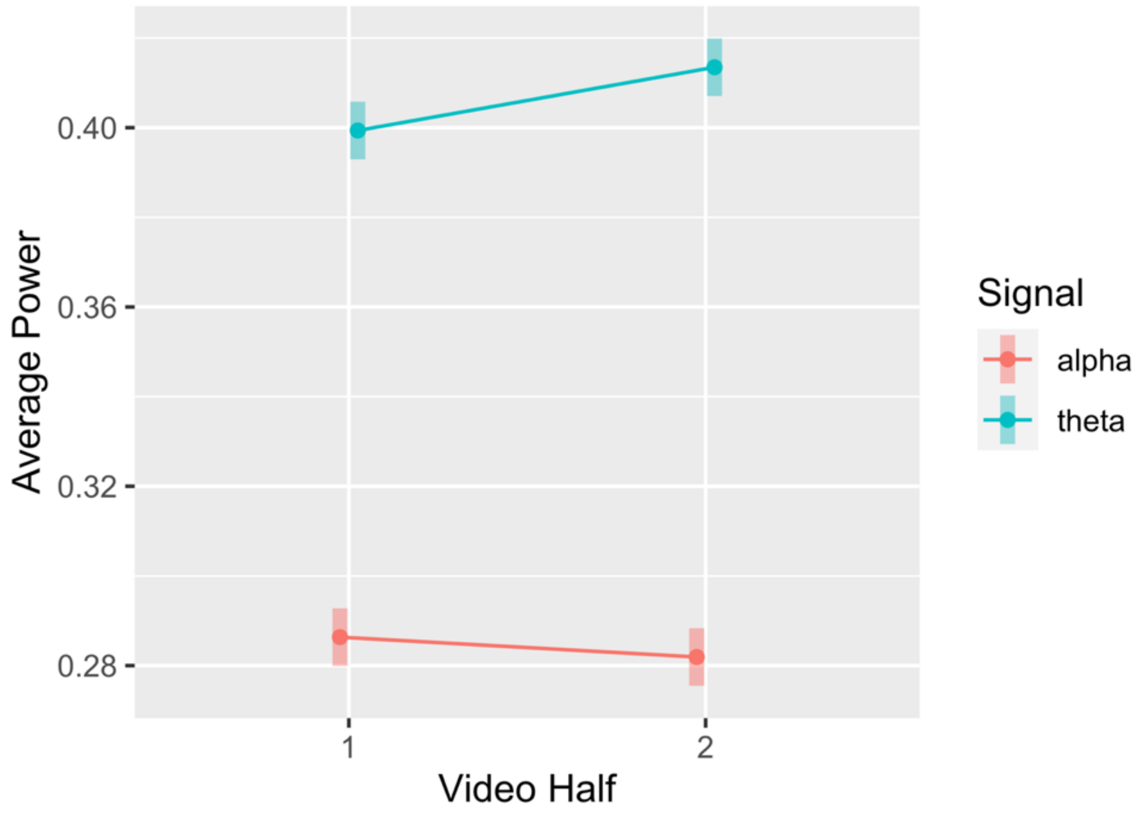


Table 3.22: Estimate, t-value, p-value, and standard errors for pairwise comparisons between relative alpha and theta power in the first and second half of video viewing in the test session. Significant effects are indicated with a *

		estimate	*t*	*p*	*SE*
Alpha, first half	Alpha, second half	0.004	0.96	.772	0.005
Theta, first half	Theta, second half	-0.014	-3.05	.013*	0.005
Alpha, first half	Theta, first half	-0.113	-24.41	< .0001*	0.005
Alpha, second half	Theta, second half	-0.132	-28.42	< .0001*	0.005

Table 3.23: Mean, standard error, and 95% confidence intervals for relative alpha and theta power in the first and second half of video viewing in the test session

Signal	*Video condition*	*M*	*SE*	95% CIs	
Alpha	First half	0.29	0.003	0.28	0.29
	Second half	0.28	0.003	0.28	0.29
Theta	First half	0.40	0.003	0.39	0.41
	Second half	0.41	0.003	0.41	0.42

3.3.2.2.2 Retest session

As in the test session, a significant main effect of video half was not found; in contrast, no significant interactions involving video half were found (Table 3.24 and Table 3.25).

Table 3.24: F-value, p-value and partial eta squared effect size for ANOVAs comparing condition effects for average relative power over the first and second half of video viewing for test and retest sessions. Significant effects are indicated with a *

	TEST			RETEST		
	F	*p*	η^2 (partial)	*F*	*p*	η^2 (partial)
Signal	1394.71	<. 0001*	0.44	2044.93	< .0001*	0.55
Region	39.59	< .0001*	0.02	31.90	< .0001*	0.02
Video number	0.72	.489	0.00	0.59	.555	
Video number	2.97	.052	0.00	7.29	.0007*	0.00
Video condition	16.91	< .0001*	0.00	17.67	< .0001*	0.01
Half	2.12	.146	0.00	1.28	.259	0.00
Signal x Region	28.89	< .0001*	0.02	33.37	< .0001*	0.02

Signal x Video number	11.44	< .0001*	0.01	1.85	.157	0.00
Region x Video number	0.32	.728	0.00	0.30	.745	0.00
Signal x Video condition	0.17	.679	0.00	3.30	.070	0.00
Region x Video condition	1.47	.226	0.00	0.88	.349	0.00
Video number x Video condition	2.11	.122	0.00	0.10	.909	0.00
Signal x Half	8.00	.005*	0.00	0.64	.423	0.00
Region x Half	0.04	.834	0.00	0.79	.373	0.00
Video number x Half	0.16	.849	0.00	0.08	.924	0.00
Video condition x Half	1.52	.218	0.00	0.01	.909	0.00
Signal x Region x Video number	0.17	.845	0.00	0.02	.978	0.00
Signal x Region x Video condition	0.05	.816	0.00	0.10	.753	0.00
Signal x Video number x Video condition	1.28	.279	0.00	0.17	.841	0.00
Region x Video number x Video condition	0.58	.560	0.00	0.05	.948	0.00
Signal x Region x Video Half	0.25	.615	0.00	0.03	.854	0.00
Signal x Video number x Half	0.18	.836	0.00	0.13	.881	0.00
Region x Video number x Half	0.08	.921	0.00	0.11	.897	0.00

Signal x Video condition x Half	0.003	.956	0.00	0.46	.500	0.00
Region x Video condition x Half	0.001	.978	0.00	0.09	.760	0.00
Video number x Video condition x Half	0.90	.409	0.00	0.16	.850	0.00
Signal x Region x Video number x Video condition	0.11	.895	0.00	0.40	.670	0.00
Signal x Region x Video number x Half	0.02	.982	0.00	0.21	.807	0.00
Signal x Region x Video condition x Half	0.04	.838	0.00	0.10	.757	0.00
Signal x Video number x Video condition x Half	0.25	.777	0.00	1.21	.299	0.00
Region x Video number x Video condition x Half	0.07	.933	0.00	0.55	.579	0.00
Signal x Region x Video number x Video condition x Half	0.08	.921	0.00	0.17	.847	0.00

Table 3.25: Mean, standard deviation and N for average relative power over video halves for test and retest sessions

Signal	Region	Video number	Video condition	Video half	*M*	*SD*	*N*	*M*	*SD*	*N*
Alpha	Frontal	First	Social	First	0.30	0.08	37	0.29	0.06	37
				Second	0.31	0.09	37	0.30	0.09	37
			Non-social	First	0.31	0.08	38	0.28	0.07	37
				Second	0.30	0.08	38	0.28	0.08	37
		Second	Social	First	0.29	0.08	37	0.27	0.08	37
				Second	0.29	0.08	37	0.27	0.08	37
			Non-social	First	0.27	0.08	38	0.27	0.07	36
				Second	0.26	0.07	38	0.26	0.07	36
		Third	Social	First	0.28	0.08	38	0.26	0.08	34
				Second	0.28	0.09	38	0.25	0.07	34
			Non-social	First	0.28	0.08	36	0.26	0.08	34
				Second	0.28	0.08	36	0.26	0.09	34
	Posterior	First	Social	First	0.30	0.06	37	0.28	0.06	37
				Second	0.30	0.07	37	0.30	0.07	37
			Non-social	First	0.30	0.06	38	0.28	0.06	37
				Second	0.29	0.06	38	0.27	0.06	37
		Second	Social	First	0.30	0.08	37	0.27	0.07	37
				Second	0.29	0.07	37	0.28	0.07	37

			Non-social	First	0.26	0.07	38	0.26	0.06	36
				Second	0.26	0.07	38	0.26	0.07	36
		Third	Social	First	0.29	0.07	38	0.27	0.07	34
				Second	0.28	0.08	38	0.27	0.06	34
			Non-social	First	0.27	0.06	36	0.25	0.06	34
				Second	0.26	0.07	36	0.26	0.06	34
Theta	Frontal	First	Social	First	0.42	0.06	37	0.46	0.07	36
				Second	0.44	0.07	37	0.45	0.08	37
			Non-social	First	0.41	0.07	38	0.43	0.07	36
				Second	0.42	0.07	38	0.45	0.06	37
		Second	Social	First	0.42	0.07	37	0.44	0.08	37
				Second	0.44	0.07	37	0.45	0.08	37
			Non-social	First	0.42	0.08	38	0.42	0.07	36
				Second	0.42	0.06	38	0.42	0.08	36
		Third	Social	First	0.43	0.06	38	0.45	0.07	34
				Second	0.43	0.06	38	0.44	0.06	34
			Non-social	First	0.43	0.07	36	0.42	0.05	34
				Second	0.44	0.07	36	0.43	0.06	34
	Posterior	First	Social	First	0.37	0.06	37	0.41	0.06	37
				Second	0.41	0.06	37	0.42	0.07	37
			Non-social	First	0.37	0.06	38	0.39	0.05	37

		Second	0.38	0.06	38	0.39	0.06	37
Second	Social	First	0.39	0.06	37	0.41	0.07	37
		Second	0.41	0.06	37	0.41	0.07	37
	Non-social	First	0.38	0.06	38	0.38	0.06	36
		Second	0.39	0.06	38	0.40	0.07	36
Third	Social	First	0.40	0.06	38	0.40	0.07	34
		Second	0.40	0.06	38	0.42	0.06	34
	Non-social	First	0.37	0.06	36	0.38	0.06	34
		Second	0.39	0.05	36	0.39	0.06	34

3.3.2.3 Power change: growth curve

To assess power change over the course of video viewing, a linear mixed model approach was used with segment number as a fixed effect and ID as a random effect allowing for random intercept and slope. Investigations revealed there were very little differences between participants' slopes, and this was not a good fit to the data, therefore the model was changed to allow random intercept only. Video number (first, second, third) and video condition (social versus non-social) were additionally included as fixed effects, so that intercepts (i.e. average power at start of video) could be compared across conditions. As average power analyses had revealed differences between theta and alpha power, and power in the frontal and posterior regions (see section 3.3.2.1), individual models were fitted for each of these.

To check whether the location of the re-reference impacted power, models were rerun including a re-reference variable as a random effect. This was a binary variable indicating whether each trial was referenced to Cz or the average of C3 and C4. For all models, reference location was found to explain only a very small amount of variance and its inclusion in the models did not impact the pattern of results for all models. Given that the location of the reference seemed to have minimal impact on EEG power, further models did not include this variable, as we chose to keep models as simple as possible. The final model (using *lmer* in R (Douglas Bates et al., 2015)) for each analysis was:

Formula: power ~ Segment number + video number + video condition + (1 | ID), data, REML = TRUE)

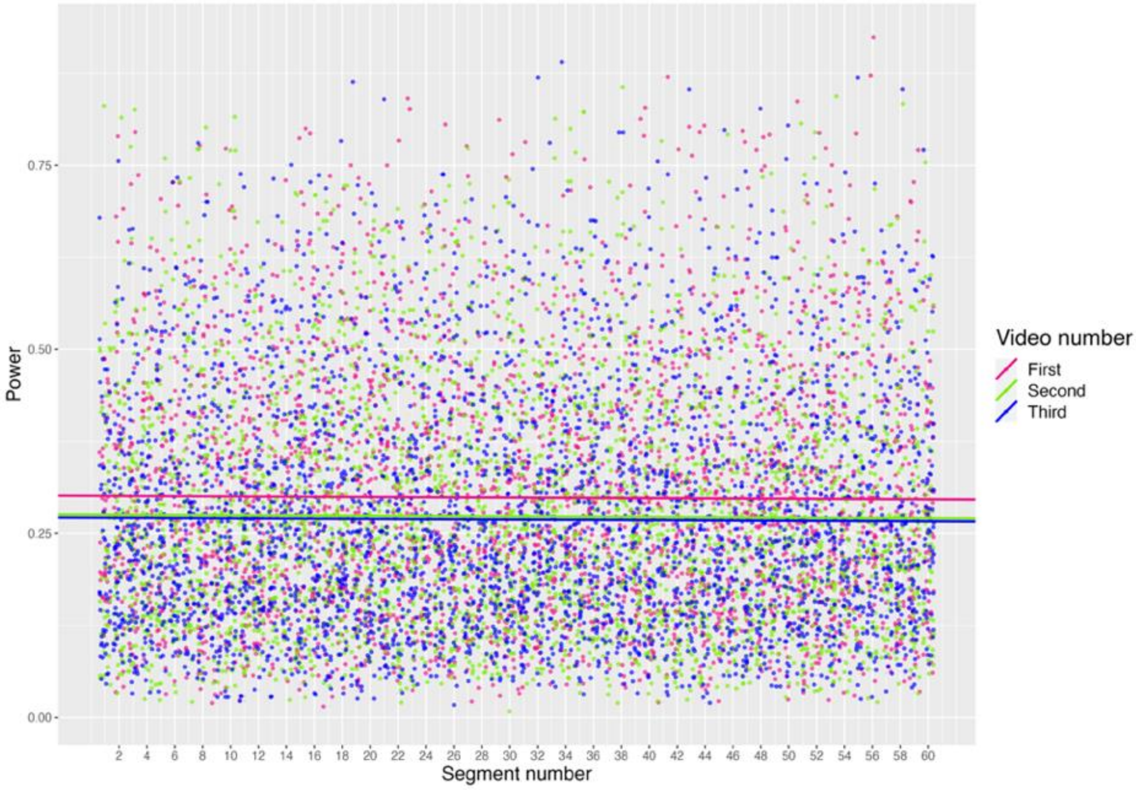
Intercept values were extracted for each participant, for use in following analyses. Any conditions where a participant provided fewer than ten usable segments of EEG were excluded and assumptions were checked. Analyses were also rerun including a minimum of 30 usable segments of EEG per conditions (see appendix B3b) and intercepts were extracted for reliability checks.

3.3.2.3.1 Test session

3.3.2.3.1.1 Frontal alpha

The intercept was significantly different from zero (β = 0.30, *SE* = 0.01, *t*(55.61) = 26.86, *p* < .001), whilst an ANOVA deviance test conducted on the model found a significant effect of video number but not condition (Table 3.26 and Table 3.28). Pairwise comparisons showed that power was higher in the first versus each of the second and third videos, with no significant differences found between the second and third videos (Table 3.27; Figure 3.19). The effect of segment number was not significant, indicating no changes in power over the course of video viewing (Table 3.28 and Table 3.29). Random effects statistics are reported in Table 3.30.

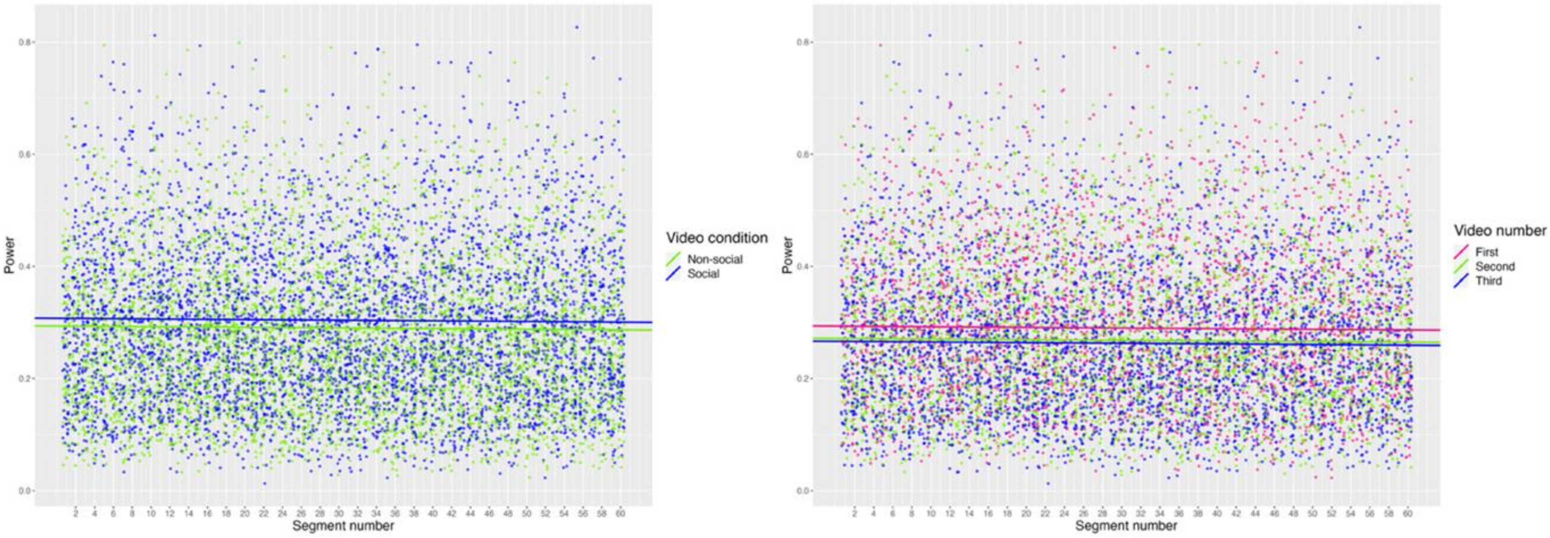
Figure 3.19: Relative frontal alpha power per segment number with lines for first, second and third estimated from linear mixed model for test session



3.3.2.3.1.2 Posterior alpha

The intercept was significantly different from zero (β = 0.29, *SE* = 0.01, *t*(54.97) = 30.35, *p* < .001), whilst an ANOVA deviance test conducted on the model found a significant effect of condition and video number (Table 3.26 and Table 3.28). Means show that power in social videos was higher than in the non-social condition, and pairwise comparisons showed that power was higher in the first versus each of the second and third videos, with no significant differences found between the second and third videos (Table 3.27; Figure 3.20). The effect of segment number was not significant, indicating no changes in power over the course of video viewing (Table 3.28 and Table 3.29). Random effects statistics are reported in Table 3.30.

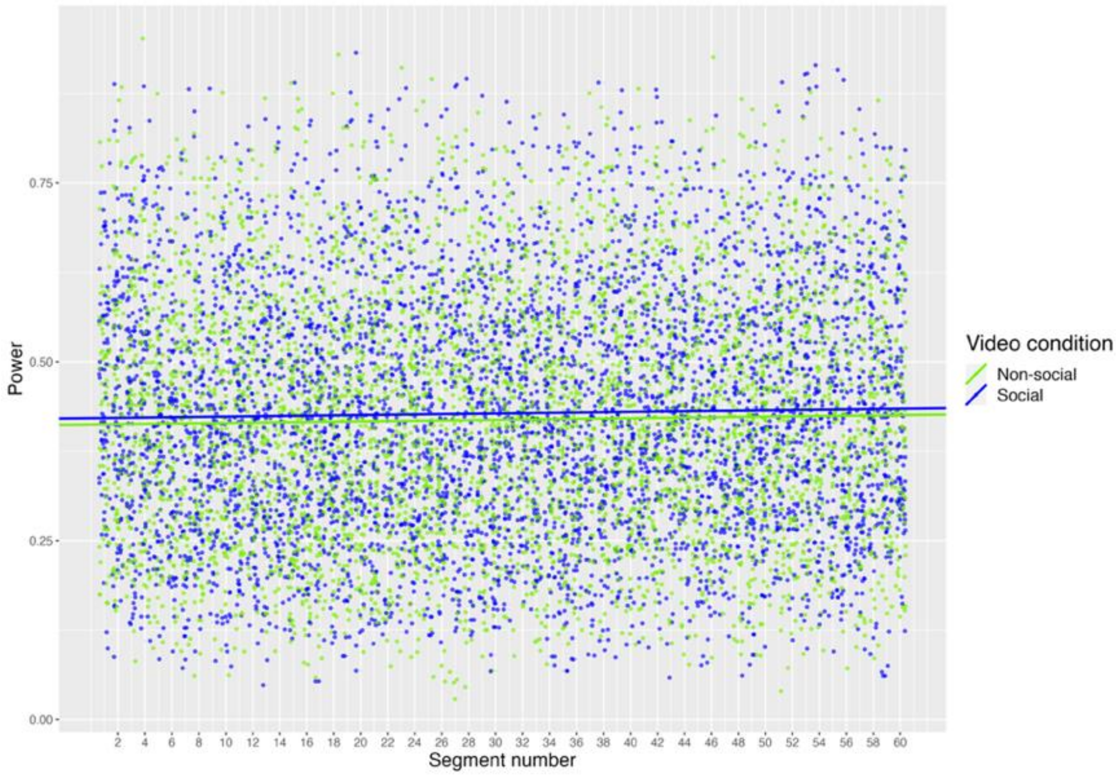
Figure 3.20: Relative posterior alpha power per segment number with lines for (a) social and non-social condition and (b) first, second and third estimated from linear mixed model for test session



3.3.2.3.1.3 Frontal theta

The intercept was significantly different from zero *(*β = 0.41*, SE* = 0.01*, t*(70.28) = 46.35, *p* < .001), whilst an ANOVA deviance test conducted on the model found a significant effect of condition and segment number but not video number (Table 3.26 and Table 3.28). Means showed power was higher in the social versus non-social condition (Figure 3.21). The estimate for segment number was small but positive, indicating that theta power increased over the course of video viewing (Table 3.28 and Table 3.29). Random effects statistics are reported in Table 3.30.

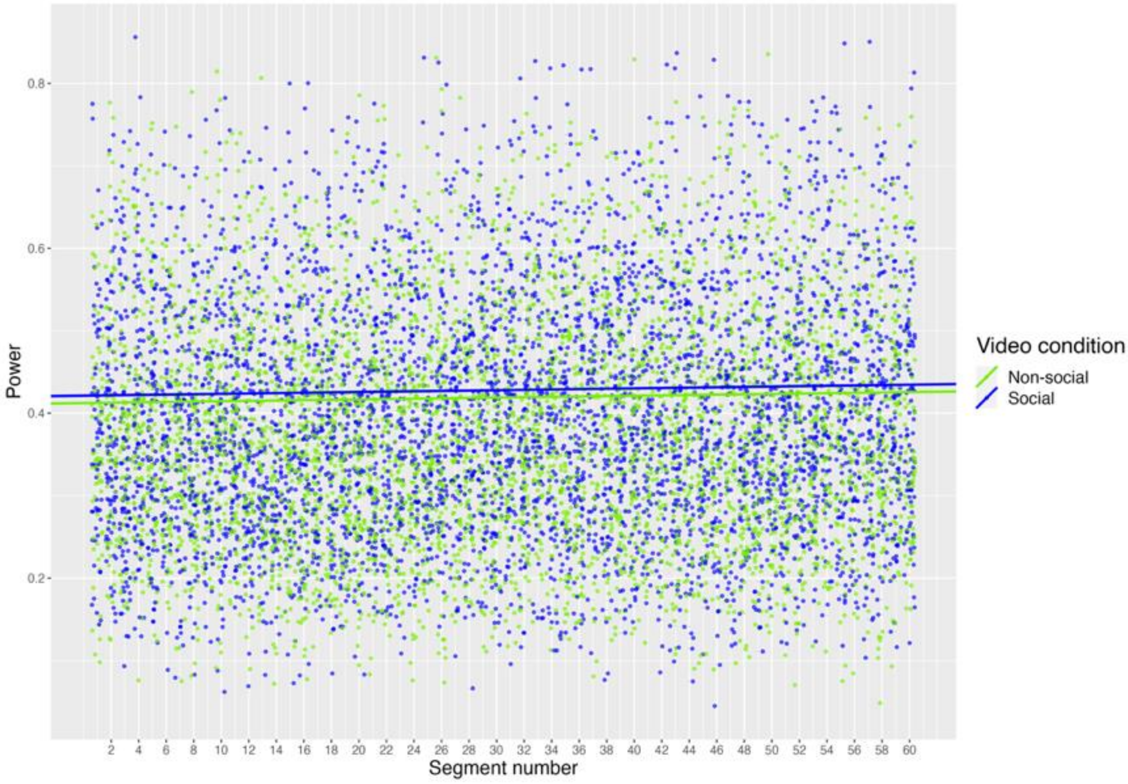
Figure 3.21: Relative frontal theta power per segment number with lines for social and non-social estimated from linear mixed model for test session



3.3.2.3.1.4 Posterior theta

The intercept was significantly different from zero (β = 0.37, *SE* = 0.01, *t*(62.96) = 44.57, *p* < .001), whilst an ANOVA deviance test conducted on the model found significant effects of condition and segment number, but not video number (Table 3.26 and Table 3.28). Power was significantly higher in the social versus non-social condition (Figure 3.22), and the estimate for segment number was small but positive, indicating that theta power increased over the course of video viewing (Table 3.28 and Table 3.29). Random effects statistics are reported in Table 3.30.

Figure 3.22: Relative posterior theta power per segment number with lines for social and non-social conditions estimated from linear mixed model for test session

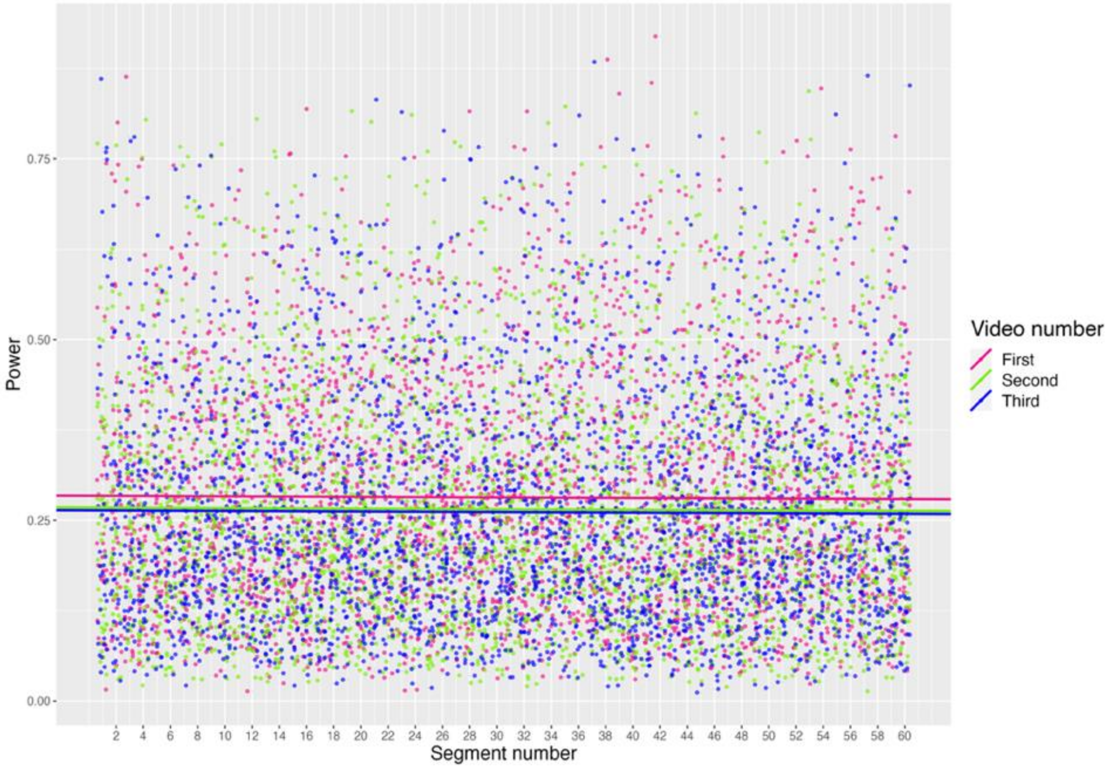


3.3.2.3.2 Retest session

3.3.2.3.2.1 Frontal alpha

The intercept was significantly different from zero (β = 0.28, *SE* = 0.01, *t*(49.21) = 25.98, $p < .0001$), whilst an ANOVA deviance test conducted on the model found a significant effect of video number but not condition (Table 3.26 and Table 3.28). Pairwise comparisons showed that power was higher in the first versus each of the second and third videos but there was no difference between first and second videos (Table 3.27; Figure 3.23). The effect of segment number was not significant, indicating no changes in power over the course of video viewing (Table 3.28 and Table 3.29). Random effects statistics are reported in Table 3.30.

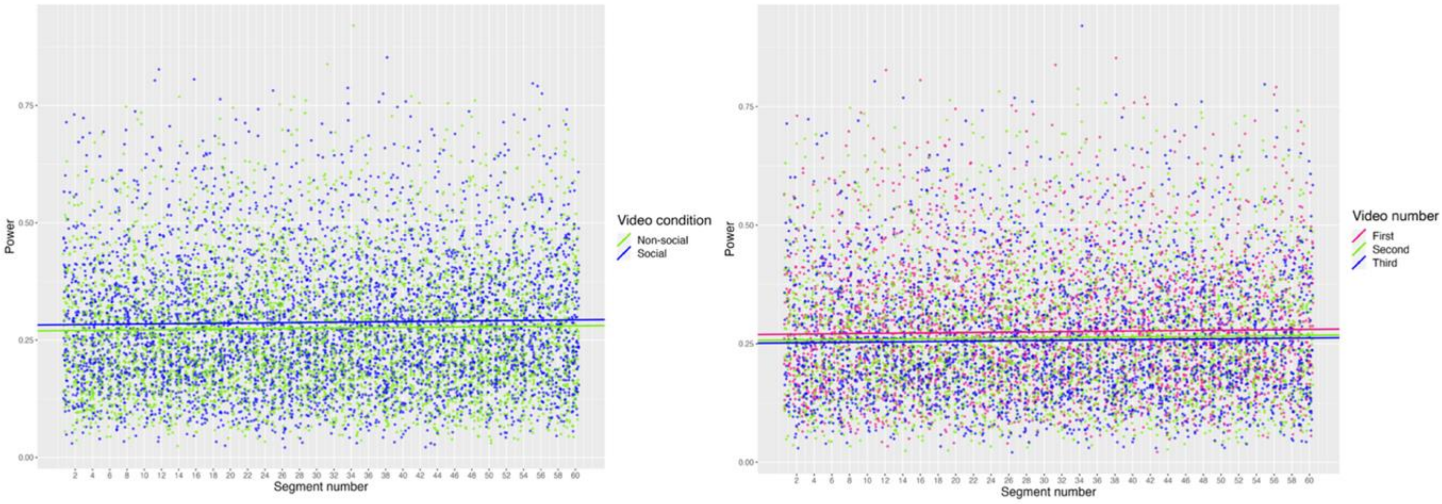
Figure 3.23: Relative frontal alpha power per segment number with lines for first, second and third estimated from linear mixed model for retest session



3.3.2.3.2.2 Posterior alpha

The intercept was significantly different from zero ($\beta = 0.27$, $SE = 0.01$, $t(49.01) = 29.20$, $p < .0001$), whilst an ANOVA deviance test conducted on the model found significant effects of segment number, condition and video number (Table 3.26 and Table 3.28). Means showed that power was lower in the non-social versus social condition, whilst pairwise comparisons showed that power was higher in the first versus second and third video but there was no difference between first and second videos (Table 3.27; Figure 3.24). The estimate for segment number was very small but positive, indicating that theta power increased by a small degree over the course of video viewing (Table 3.28 and Table 3.29). Random effects statistics are reported in Table 3.30.

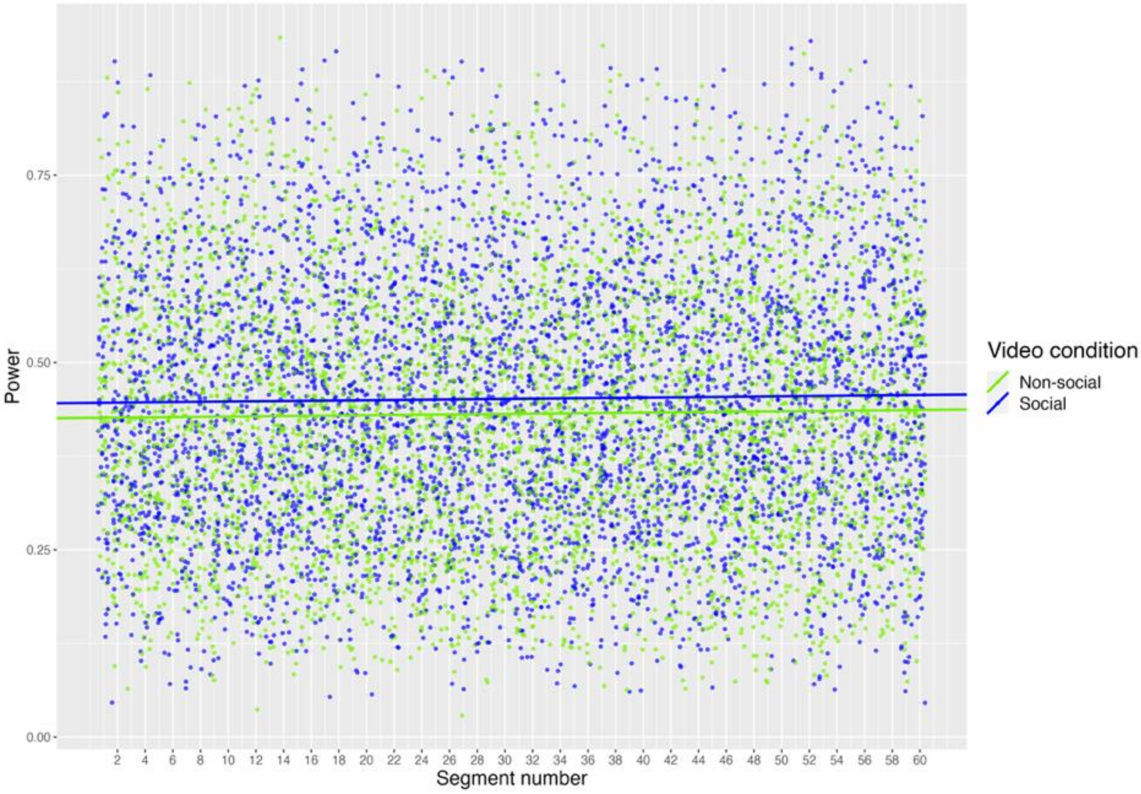
Figure 3.24: Relative posterior alpha power per segment number with lines for (a) social and non-social condition and (b) first, second and third estimated from linear mixed model for retest session



3.3.2.3.2.3 Frontal theta

The intercept was significantly different from zero *(*β = 0.43*, SE* = 0.01*, t*(56.03) = 42.39, *p* < .0001), whilst an ANOVA deviance test conducted on the model found a significant effect of condition but not segment number or video repetitions (Table 3.26 and Table 3.28). Means showed that power was lower in the non-social versus social condition (Figure 3.25). The estimate for segment number was not significant, indicating no changes in power over the course of video viewing (Table 3.28 and Table 3.29). Random effects statistics are reported in Table 3.30.

Figure 3.25: Relative frontal theta power per segment number with lines for social and non-social estimated from linear mixed model for retest session



3.3.2.3.2.4 Posterior theta

The intercept was significantly different from zero (β = 0.38, *SE* = 0.01, *t*(51.62) = 41.62, *p* < .0001), whilst an ANOVA deviance test conducted on the model found a significant effect of segment number and video condition, but not video number (Table 3.26 and Table 3.28). Means showed that power was higher in the social versus non-social condition (Table 3.26; Figure 3.26). The estimate for segment number was very small but positive, indicating that theta power increased by a small degree over the course of video viewing (Table 3.28 and Table 3.29). Random effects statistics are reported in Table 3.30.

Figure 3.26: Relative posterior alpha theta per segment number with lines for social and non-social conditions estimated from linear mixed model for retest session

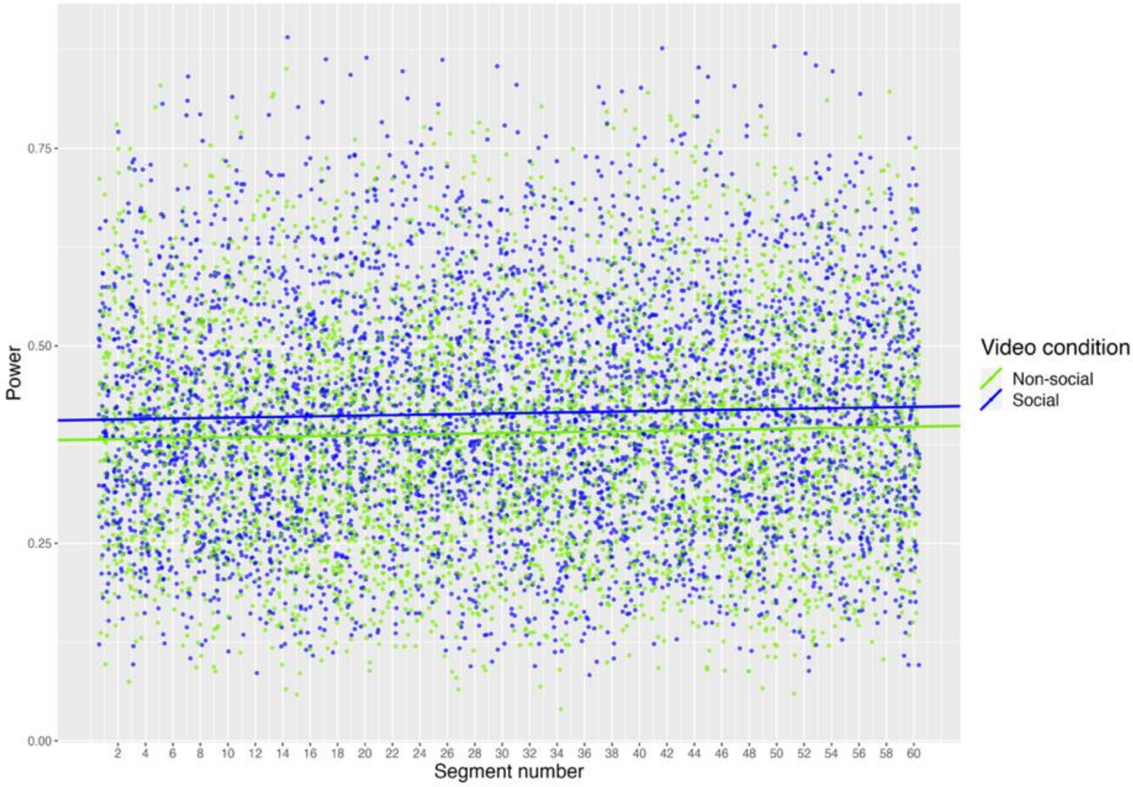


Table 3.26: Mean, standard error and 95% confidence intervals for first, second and third videos, and social and non-social conditions from average relative power per segment in the test and retest session. Statistics are shown for each of frontal and posterior theta and alpha power

		TEST				RETEST			
		Mean	*SE*	95% CIs		*Mean*	*SE*	95% CIs	
				Video number					
	First	0.30	0.01	0.26	0.30	0.28	0.01	0.26	0.31
	Second	0.28	0.01	0.25	0.30	0.27	0.01	0.25	0.29
Frontal alpha	Third	0.27	0.01	0.25	0.29	0.26	0.01	0.24	0.29
				Video condition					
	Social	0.29	0.01	0.27	0.31	0.27	0.01	0.25	0.29
	Non-social	0.28	0.01	0.26	0.30	0.27	0.01	0.25	0.29
				Video number					
	First	0.30	0.01	0.28	0.32	0.28	0.01	0.26	0.30
	Second	0.28	0.01	0.25	0.30	0.27	0.01	0.25	0.29
Posterior alpha	Third	0.27	0.01	0.25	0.29	0.26	0.01	0.25	0.28
				Video condition					
	Social	0.29	0.01	0.27	0.31	0.28	0.01	0.26	0.30
	Non-social	0.27	0.01	0.26	0.29	0.27	0.01	0.25	0.28
				Video number					
	First	0.42	0.01	0.41	0.44	0.44	0.01	0.42	0.46
Frontal theta	Second	0.43	0.01	0.41	0.44	0.44	0.01	0.42	0.45
	Third	0.43	0.01	0.41	0.45	0.43	0.01	0.41	0.45
				Video condition					

	Social	0.43	0.01	0.41	0.44	0.45	0.01	0.43	0.47
	Non-social	0.42	0.01	0.41	0.44	0.43	0.01	0.41	0.45
				Video number					
	First	0.39	0.01	0.37	0.40	0.40	0.01	0.39	0.42
	Second	0.39	0.01	0.38	0.41	0.40	0.01	0.38	0.42
Posterior theta	Third	0.39	0.01	0.37	0.40	0.40	0.01	0.38	0.42
				Video condition					
	Social	0.40	0.08	0.38	0.41	0.41	0.01	0.40	0.43
	Non-social	0.38	0.08	0.36	0.39	0.39	0.01	0.37	0.41

Table 3.27: Estimate, standard error, t-value, and p-value for pairwise comparisons of average relative power per segment number between video numbers for frontal and posterior alpha in test and retest conditions. Significant effects are indicated with a *

			TEST				RETEST			
			β	*SE*	*t*	*p*	β	*SE*	*t*	*p*
Frontal alpha	First	Second	0.03	0.004	6.93	< .0001*	0.02	0.004	4.41	< .0001*
	First	Third	0.03	0.004	8.08	< .0001*	0.02	0.004	5.43	< .0001*
	Second	Third	0.004	0.004	1.15	.76	0.004	0.004	1.07	.85
Posterior alpha	First	Second	0.02	0.003	7.02	< .0001*	0.01	0.003	4.06	.0001*
	First	Third	0.03	0.003	8.78	< .0001*	0.02	0.003	5.82	< .0001*
	Second	Third	0.01	0.003	1.77	.23	0.01	0.003	1.81	.21

Table 3.28: Chi-squared, p-value and partial eta squared effect size of effects of segment number, video number and condition from ANOVAs on linear mixed models of average relative power for each of frontal and posterior regions, and alpha and theta bands, for the test and retest sessions

		TEST			RETEST		
		$\chi 2$	*p*	η^2 (partial)	$\chi 2$	*p*	η^2 (partial)
Frontal alpha	Segment number	0.89	.345	<.001	0.79	.38	<.001
	Video number	75.30	< .0001*	0.007	33.33	< .0001*	0.002
	Condition	3.86	.050	<.001	1.54	.22	<.001
Posterior alpha	Segment number	2.37	.12	<.001	5.40	.02*	<.001
	Video number	85.17	< .0001*	0.008	35.63	< .0001*	0.001
	Condition	30.42	< .0001*	0.003	24.86	< .0001*	0.002
Frontal theta	Segment number	5.77	.02*	<.001	3.28	.07	<.001
	Video number	3.71	.16	<.001	4.58	.10	<.001
	Condition	7.78	.005*	<.001	35.97	< .0001*	0.001
Posterior theta	Segment number	26.79	< .0001*	0.003	11.78	.0006*	<.001
	Video number	1.18	.55	<.001	2.19	.33	<.001
	Condition	63.72	< .0001*	0.006	83.38	< .0001*	0.006

Table 3.29: Estimate, standard error, degrees of freedom and 95% confidence intervals for each of the fixed effects in linear mixed models performed on average relative power per segment for each of frontal and posterior alpha and theta in the test and retest sessions

	Fixed Effects									
	TEST					RETEST				
	β	*SE*	*DF*	CIs 95%		*β*	*SE*	*DF*	CIs 95%	
Frontal alpha										
Intercept	0.30	0.01	55.61	0.29	0.33	0.28	0.01	49.21	0.27	0.30
Segment Number	<-.001	<.001	10096.28	<-.001	<.001	<-.001	<.001	9309.27	<-.001	<.001
Video number: second	-0.03	0.004	10124.68	-0.03	-0.02	-0.02	0.004	9321.88	-0.02	-0.01
Video number: third	-0.03	0.004	10130.20	-0.04	-0.02	-0.02	0.004	9319.62	-0.03	-0.02
Video condition: social	0.01	0.003	10126.33	<0.001	0.01	0.004	0.003	9317.36	<.001	0.01
Posterior alpha										
Intercept	0.29	0.01	54.97	0.28	0.31	0.27	0.01	49.01	0.26	0.28
Segment Number	<-.001	<.001	10096.25	<-.001	<.001	<0.001	<0.001	9309.10	<-.001	<.001
Video number: second	-0.02	0.003	10123.51	-0.03	-0.01	-0.01	0.003	9321.79	-0.02	-0.01
Video number: third	-0.03	0.003	10129.02	-0.03	-0.02	-0.02	0.003	9319.52	-0.02	-0.02
Video condition: social	0.01	0.003	10125.15	0.01	0.02	0.01	0.003	9317.25	0.010	0.017
Frontal theta										
Intercept	0.41	0.01	70.28	0.40	0.43	0.43	0.01	56.03	0.41	0.44
Segment Number	<.001	<0.001	10098.13	<0.001	<0.001	<0.001	<0.001	9309.86	<-0.001	<0.001

Video number: second	0.001	0.004	10137.87	-0.003	0.013	-0.01	0.004	9328.10	-0.012	0.002
Video number: third	0.007	0.004	10138.85	-0.001	0.013	-0.01	0.004	9325.13	-0.017	-0.006
Video condition: social	0.009	0.003	10138.52	0.003	0.015	0.02	0.003	9321.85	0.016	0.026
Posterior theta										
Intercept	0.37	0.01	62.96	0.36	0.38	0.38	0.01	51.62	0.37	0.39
Segment Number	<.001	<0.001	10097.23	<0.001	0.001	<.001	<0.001	9309.38	<0.001	0.001
Video number: second	0.003	0.003	10133.58	-0.001	0.013	-0.002	0.003	9324.33	-0.006	0.002
Video number: third	0.001	0.003	10137.66	-0.006	0.006	-0.005	0.003	9321.74	-0.012	-0.003
Video condition: social	0.021	0.003	10134.92	0.016	0.026	0.025	0.003	9319.06	0.022	0.030

Table 3.30: Estimate and standard error for each of the random effects in linear mixed models performed on average relative power per segment for each of frontal and posterior alpha and theta in the test and retest sessions

		Random Effects			
		TEST		RETEST	
		β	*SE*	β	*SE*
Frontal alpha	ID: Intercept	0.005	0.07	0.004	0.06
	Residual	0.022	0.15	0.021	0.14
Posterior alpha	ID: Intercept	0.004	0.06	0.003	0.05
	Residual	0.015	0.01	0.015	0.12
Frontal theta	ID: Intercept	0.003	0.05	0.003	0.06
	Residual	0.03	0.16	0.017	0.16
Posterior theta	ID: Intercept	0.002	0.05	0.003	0.05
	Residual	0.017	0.13	0.017	0.01

3.3.2.4 Variability

Table 3.31 shows descriptive data for variability measures per signal (alpha and theta) and region (frontal and posterior). Residual standard errors were calculated in relation to regression lines estimated by the linear mixed models described in section 3.3.2.3. and standard error of estimates were calculated using the *se.ranef* function in R (Douglas Bates et al., 2015). Standard deviations were calculated in relation to mean EEG power when averaged over the whole video; these are closely related to residual standard errors.

Table 3.31: Mean, standard deviation and number of participants for residual standard error, standard error of estimate and standard deviation of relative alpha and theta power in frontal and posterior regions

Variability statistic			TEST			RETEST		
	Signal	*Region*	*M*	*SD*	*N*	*M*	*SD*	*N*
Residual standard error	Alpha	Frontal	0.02	0.01	45	0.02	0.01	39
		Posterior	0.02	0.01		0.02	0.01	
	Theta	Frontal	0.03	0.004		0.03	0.004	
		Posterior	0.02	0.003		0.02	0.003	
Standard error of estimates	Alpha	Frontal	0.01	0.003	45	0.01	0.001	39
		Posterior	0.01	0.002		0.01	0.002	
	Theta	Frontal	0.01	0.003		0.01	0.001	
		Posterior	0.01	0.002		0.01	0.001	
Standard deviation	Alpha	Social	0.14	0.03	45	0.14	0.02	39
		Non-social	0.12	0.02		0.12	0.02	
	Theta	Social	0.16	0.01		0.16	0.01	
		Non-social	0.13	0.01		0.13	0.01	

3.3.3 Test-retest reliability

Reliability analyses were conducted using a range of measures and across a range of conditions; ICC statistics are shown in Table 3.32. Measures of power averaged over the whole head showed mostly fair and good reliability; this was particularly so for theta power. Differences between average power over the whole video in the frontal and posterior regions showed some fair and some poor reliability, with more fair values for alpha, and more poor values for theta power. Differences in power between the first and second half of video viewing showed mostly poor ICC values. Intercept values from linear mixed models had mostly fair and good reliability, with particularly good ICC values for relative theta power over the whole head and each of the frontal and posterior regions. ICC values for variability measures were generally poor, with fair reliability for some alpha measures. As previous work had indicated that a minimum of 20 segments were needed for good reliability and 40 for excellent reliability (Troller-Renfree et al., 2021), reliability analyses were rerun using a minimum of 30 segments. Results showed a similar pattern of reliability to those requiring a smaller number of good epochs; see Table B 28. These cut-offs were not used for main analyses as they lead to small sample sizes, but were included as robustness checks.

Table 3.32: ICC statistics for reliability analyses between test and retest sessions for a number of measures. Statistics are coloured according to ICC values: values below 0.40 are red, values from 0.40 to 0.75 are orange, and values above 0.75 are green

Measure	Condition	Absolute Alpha ICC	LB	UB	*p*	*n*	Absolute Theta ICC	LB	UB	*p*	*n*	Relative Alpha ICC	LB	UB	*p*	*n*	Relative Theta ICC	LB	UB	*p*	*n*
	Average across regions																				
	Social, first	.343	-0.03	0.63	.036	27	.634	0.34	0.82	.000	27	.534	0.20	0.76	.002	27	.728	0.49	0.87	.000	27
	Social, second	.500	0.18	0.73	.002	30	.578	0.28	0.77	.000	30	.539	0.24	0.75	.001	30	.772	0.57	0.88	.000	30
	Social, third	.451	0.11	0.70	.006	29	.630	0.35	0.81	.000	29	.634	0.36	0.81	.000	29	.755	0.54	0.88	.000	29
	Nonsocial, first	.601	0.31	0.79	.000	30	.689	0.44	0.84	.000	30	.713	0.43	0.86	.000	30	.775	0.56	0.89	.000	30
	Nonsocial, second	.625	0.34	0.80	.000	29	.707	0.46	0.85	.000	29	.753	0.54	0.88	.000	29	.799	0.61	0.90	.000	29
	Nonsocial, third	.514	0.16	0.75	.004	26	.748	0.52	0.88	.000	26	.596	0.28	0.80	.001	26	.678	0.40	0.84	.000	26
	Social	.244	-0.28	0.65	.174	16	.763	0.45	0.91	.000	16	.701	0.34	0.88	.001	16	.886	0.70	0.96	.000	16
	Nonsocial	.531	0.08	0.80	.012	16	.797	0.51	0.92	.000	16	.702	0.35	0.88	.001	16	.908	0.76	0.97	.000	16
	First	.391	0.02	0.67	.020	25	.740	0.49	0.88	.000	25	.609	0.30	0.81	.000	25	.801	0.56	0.91	.000	25
	Second	.577	0.26	0.78	.001	27	.682	0.41	0.84	.000	27	.715	0.47	0.86	.000	27	.816	0.63	0.91	.000	27
Average over whole video	Third	.407	0.00	0.70	.026	22	.822	0.62	0.92	.000	22	.613	0.28	0.82	.001	22	.778	0.54	0.90	.000	22
	Difference between regions																				
	Social, first	.534	0.19	0.76	.002	27	.372	-0.01	0.66	.028	27	.304	-0.07	0.61	**.056**	27	.651	0.36	0.82	.000	27
	Social, second	.299	-0.07	0.59	.055	30	.305	-0.07	0.60	**.052**	30	.479	0.14	0.71	.004	30	.201	-0.18	0.52	**.144**	30
	Social, third	.536	0.21	0.75	.001	29	.257	-0.12	0.57	**.090**	29	.313	-0.06	0.61	.049	29	.108	-0.27	0.45	**.289**	29
	Nonsocial, first	.443	0.10	0.69	.007	30	.173	-0.18	0.49	**.168**	30	.111	-0.24	0.44	**.269**	30	-.024	-0.38	0.33	**.550**	30
	Nonsocial, second	.508	0.17	0.74	.002	29	.343	-0.03	0.63	.035	29	.512	0.18	0.74	.002	29	.279	-0.10	0.58	**.072**	29
	Nonsocial, third	-.021	-0.41	0.37	.541	25	-.063	-0.45	0.34	**.617**	25	.017	-0.30	0.37	**.460**	25	-.257	-0.61	0.16	**.886**	25
	Social	.494	0.01	0.79	.023	16	.311	-0.23	0.70	**.124**	16	.558	0.11	0.82	.009	16	.490	0.00	0.79	.025	16
	Nonsocial	.549	0.09	0.81	.011	16	.349	-0.18	0.72	**.093**	16	.666	0.29	0.87	.001	16	.252	-0.30	0.66	**.176**	16
	First	.563	0.22	0.78	.002	25	.308	-0.09	0.62	**.062**	25	.157	-0.21	0.50	**.201**	25	.448	0.07	0.71	.011	25
	Second	.416	0.04	0.69	.016	27	.345	-0.04	0.64	.040	27	.685	0.42	0.84	.000	27	.241	-0.16	0.57	**.115**	27
Average over whole video	Third	.591	0.23	0.81	.002	22	.351	-0.09	0.67	**.056**	22	.284	-0.16	0.63	**.100**	22	.224	-0.22	0.59	**.157**	22
	Average across regions																				
	Social, first	-.517	-0.78	-0.15	**.996**	27	.223	-0.10	0.53	**.091**	27	-.439	-0.73	-0.05	**.987**	27	-.106	-0.37	0.22	**.751**	27
	Social, second	.276	-0.11	0.58	**.075**	29	.150	-0.15	0.45	**.173**	29	.116	-0.27	0.46	**.275**	29	-.064	-0.36	0.27	**.649**	29
	Social, third	.109	-0.28	0.46	**.292**	28	.040	-0.34	0.41	**.419**	28	-.151	-0.51	0.24	**.774**	28	.141	-0.25	0.49	**.236**	28
	Nonsocial, first	.079	-0.27	0.42	**.332**	30	-.284	-0.59	0.09	**.934**	30	-.029	-0.38	0.33	**.561**	30	-.029	-0.39	0.34	**.559**	30
	Nonsocial, second	-.311	-0.63	0.08	**.942**	28	.092	-0.29	0.45	**.321**	28	.054	-0.34	0.42	**.394**	28	-.329	-0.65	0.06	**.951**	28
	Nonsocial, third	-.173	-0.55	0.25	**.791**	25	.161	-0.23	0.51	**.208**	25	-.224	-0.56	0.18	**.867**	25	-.082	-0.45	0.31	**.658**	25
	Social	-.090	-0.61	0.44	.622	15	.109	-0.24	0.52	**.291**	15	-.165	-0.66	0.38	**.719**	15	-.271	-0.52	0.23	**.891**	15
	Nonsocial	-.280	-0.74	0.29	.834	15	-.238	-0.71	0.32	**.797**	15	-.527	-0.89	0.03	**.969**	15	-.395	-0.82	0.18	**.914**	15
	First	-.122	-0.49	0.28	.724	25	.047	-0.36	0.43	**.411**	25	-.421	-0.71	-0.02	**.981**	25	-.255	-0.57	0.14	**.903**	25
	Second	-.208	-0.58	0.21	.834	25	.190	-0.14	0.51	**.134**	25	.048	-0.36	0.44	**.410**	25	-.261	-0.57	0.13	**.910**	25
Difference between halves of video	Third	.267	-0.19	0.62	.123	21	.154	-0.31	0.55	**.253**	21	.230	-0.16	0.58	**.127**	21	-.226	-0.61	0.23	**.837**	21
	Difference between regions																				
	Social, first	.452	0.10	0.70	.006	27	.182	-0.21	0.52	**.181**	27	.282	-0.11	0.60	**.078**	27	.434	0.07	0.70	.011	27
	Social, second	.292	-0.09	0.59	.064	29	-.158	-0.48	0.21	**.804**	29	-.089	-0.45	0.29	**.676**	29	.018	-0.31	0.36	**.459**	29
	Social, third	-.200	-0.55	0.19	.840	28	.147	-0.23	0.49	**.223**	28	-.007	-0.39	0.37	**.513**	28	.425	0.06	0.69	.012	28
	Nonsocial, first	-.354	-0.65	0.02	.969	30	-.055	-0.40	0.30	**.618**	30	-.288	-0.61	0.09	**.934**	30	-.019	-0.38	0.34	**.540**	30
	Nonsocial, second	-.205	-0.51	0.16	.870	28	-.065	-0.44	0.32	**.626**	28	.298	-0.08	0.60	**.058**	28	-.087	-0.46	0.30	**.669**	28
	Nonsocial, third	-.021	-0.41	0.37	.541	25	-.063	-0.45	0.34	**.617**	25	.017	-0.30	0.37	**.460**	25	-.257	-0.61	0.16	**.886**	25
	Social	.313	-0.20	0.70	.112	15	.344	-0.20	0.72	**.102**	15	.314	-0.14	0.69	**.089**	15	.486	0.00	0.79	.026	15
	Nonsocial	.039	-0.43	0.51	.440	15	-.188	-0.68	0.36	**.745**	15	-.297	-0.70	0.24	**.869**	15	-.079	-0.61	0.46	**.607**	15
	First	-.009	-0.39	0.38	.518	25	.077	-0.32	0.45	**.354**	25	-.287	-0.63	0.13	**.914**	25	.387	0.01	0.67	.023	25
	Second	.300	-0.10	0.62	.068	25	-.273	-0.63	0.15	**.900**	25	.088	-0.32	0.46	**.337**	25	-.148	-0.53	0.27	**.756**	25
Difference between halves of video	Third	-.101	-0.54	0.35	.664	21	.138	-0.33	0.54	**.278**	21	.051	-0.37	0.46	**.410**	21	.312	-0.15	0.65	**.086**	21
	LMER																				
	Differences	.584	0.32	0.77	.000	35	.355	0.03	0.61	.018	35	.476	0.18	0.69	.001	35	.270	-0.07	0.55	**.060**	35
	Wholehead	.527	0.25	0.73	.000	35	.754	0.57	0.87	.000	35	.681	0.41	0.83	.000	35	.815	0.65	0.90	.000	35
	Frontal	.636	0.39	0.80	.000	35	.667	0.44	0.82	.000	35	.683	0.45	0.83	.000	35	.752	0.56	0.87	.000	35
Intercepts from LMERs	Posterior	.432	0.13	0.66	.003	35	.680	0.45	0.82	.000	35	.665	0.35	0.83	.000	35	.797	0.63	0.89	.000	35
	Variability																				
	Differences	-.033	-0.33	0.28	.582	35	.107	-0.23	0.42	**.267**	35	.384	0.06	0.64	.012	35	-.241	-0.54	0.10	**.915**	35
	Wholehead	.420	0.10	0.66	.006	35	.271	-0.07	0.55	**.056**	35	.438	0.12	0.67	.004	35	.390	0.07	0.64	.010	35
	Frontal	.252	-0.07	0.53	.063	35	.371	0.06	0.62	.011	35	.480	0.18	0.70	.002	35	.187	-0.15	0.49	**.139**	35
RSE	Posterior	.328	0.00	0.59	.024	35	-.023	-0.36	0.32	**.552**	35	.370	0.04	0.62	.014	35	.268	-0.08	0.55	**.061**	35
	Differences	-.058	-0.35	0.26	.642	35	.004	-0.33	0.33	**.491**	35	.332	0.00	0.60	.026	35	-.285	-0.58	0.06	**.949**	35
	Wholehead	.418	0.10	0.66	.006	35	.286	-0.05	0.56	.046	35	.487	0.19	0.70	.001	35	.442	0.13	0.67	.004	35
	Frontal	.277	-0.05	0.55	.048	35	.384	0.07	0.63	.009	35	.533	0.25	0.73	.000	35	.219	-0.12	0.51	**.100**	35
SD	Posterior	.306	-0.02	0.57	.032	35	.007	-0.34	0.34	**.485**	35	.413	0.10	0.65	.006	35	.317	-0.02	0.59	.032	35
	Differences	.221	-0.08	0.50	.073	35	.272	-0.05	0.55	.049	35	.192	-0.11	0.48	**.109**	35	.251	-0.09	0.54	**.071**	35
	Wholehead	.205	-0.12	0.50	.110	35	.213	-0.12	0.51	**.105**	35	.204	-0.12	0.49	**.107**	35	.224	-0.12	0.52	**.095**	35
	Frontal	.213	-0.12	0.51	.106	35	.220	-0.12	0.51	**.100**	35	.203	-0.12	0.49	**.107**	35	.227	-0.11	0.52	**.093**	35
SE_Ranef	Posterior	.199	-0.12	0.49	.113	35	.207	-0.13	0.50	**.110**	35	.205	-0.12	0.50	**.108**	35	.222	-0.12	0.51	**.097**	35

3.4 DISCUSSION

This study aimed to investigate condition comparisons and assessed the reliability of numerous EEG power measures using a low-density, portable EEG system with 2.5- to 4-year-old children.

3.4.1 Condition comparisons

Condition comparisons revealed significant differences in EEG power depending on factors such as frequency band, brain region and video condition. Results showed that theta power was greater than alpha power, which was found consistently for relative and absolute power, and for whole and half videos. That theta power was consistently higher than alpha power may give some insight into the relative contributions of these two frequency domains in toddler-aged children. Previous work suggests a shift from theta to alpha as the dominant frequency in the early years of life, though existing literature is not clear when this shift occurs. The current work might suggest that alpha dominance occurs after 4 years of age, which lends support to Cellier et al. (2021) who suggest this shift occurs at around 7 years old.

There was also a tendency for greater power in the posterior versus frontal region for absolute power, though this pattern appeared reversed for relative power. This seemed particularly driven by regional differences in relative theta power, whereas relative alpha power was similar across regions. This finding is somewhat puzzling when considered in the context of work which indicates a relation between posterior alpha power and anticipatory attention (Orekhova et al., 2001). In that work, findings are similar for both relative and z-scored absolute power, though absolute power as used here was not analysed in this way. Notably, the current study used a passive video viewing design, whilst the increase in posterior alpha was found during the anticipatory game of a peek-a-boo (Orekhova et al., 2001). It is therefore possible that these designs are tapping different attentional skills, with Orekhova et al. (2001) suggesting posterior alpha may be involved in inhibition of attention shifting networks, which may be unlikely to be involved in video viewing in the current study. Whereas previous work has implicated posterior alpha in maintaining focussed attention to the centre of the visual field, the videos in the current study did not demand such specific focus and may be the cause of differences between these studies. In addition, the peek-a-boo design was a live interaction with a researcher, whereas the current study utilised a video which participants passively watched; this difference could also lead to discrepancy in processing and neural engagement. It is also possible that differences in findings are due to developmental differences, given the different ages of participants in the two studies. Nonetheless, absolute power did show the predicted higher alpha power in the posterior versus frontal region, which appears in line with work suggesting a role of posterior alpha in maintaining centrally focussed attention.

The finding that theta power was greater in posterior versus frontal regions is in contrast to findings of maximal theta power at frontal areas (Canen & Brooker, 2017). However, greater posterior theta power here does support a potential shift from frontal-specific to more widespread theta power from infancy to early childhood (Cuevas et al., 2012; Orekhova et al., 2006). Indeed, some existing research indicated predominantly occipital/ posterior theta power specifically during memory-encoding (Cuevas et al., 2012) and social processing (Orekhova et al., 2006). When considered with the current findings, this may support a change of dominance in theta power from frontal to posterior regions in toddlerhood and preschool age.

As was hypothesised, EEG power was consistently higher for the social compared to non-social video condition for relative and absolute, as well as alpha and theta power. This was in line with previous work in younger children (Jones et al., 2015) and extends this finding to this older age-range. That alpha power was larger in the social condition also supports a role for alpha in information processing; in particular, given that social interaction is thought to provide a greater amount of information to process, it supports a role for alpha in memory encoding (Cuevas et al., 2012). Findings in the current age range are generally lacking, therefore the finding that social conditions induce more (absolute and relative) power in both alpha and theta bands is a valuable contribution to existing work.

As has previously been found in infants (Braithwaite et al., 2019; Jones et al., 2020), the current study found some evidence of an increase in theta power over the course of video viewing, thereby also extending this finding to toddlers. Of note, this finding was specific to relative theta but not alpha power when comparing average power in the second compared to first 30 seconds of the video. Absolute power was higher in the second half of videos when collapsed across both theta and alpha power only, again indicating a change in processing over time in relation to more general EEG power. Continuous measures of average power also found a small but significant increase over the course of video viewing for some measures. This was especially the case for theta measures, with only relative frontal theta in the retest not finding an increase. By contrast, only absolute frontal alpha in the test and relative posterior alpha in the retest sessions similarly found a significant change. These findings support that specifically theta may be involved with cognitive processing throughout video viewing, though the relatively small degree of increase along with some null findings may suggest a less robust finding than has been found in infants. There was additionally no evidence of any differences in the degree of theta power change across video conditions, which differed from what was hypothesised, though this finding was based upon findings in much younger infants (Jones et al., 2020) and may be less apparent in older children.

Of note, higher relative theta power in the second versus first half of video viewing was only found in the test but not retest session, which might indicate a relation to repeated exposure or learning from visit one to visit two. Repetitions of videos within a session, however, showed no differences in theta power, thereby suggesting the interval between visit one and two led to these differences. If sample sizes provided enough statistical power, it might have been fruitful to investigate whether differences in the degree of change relate to the length of interval between sessions.

The lack of difference in theta power between video repetitions was contrary to a hypothesised increase over subsequent viewings, as predicted based on Meyer et al. (2019). In that study, preschoolers completed various cognitive tasks; it was found that EEG power was greater during presentation of a fixation cross which preceded a task in relation to increased task engagement. Here, there was no task per se for participants to complete and the study design may not have engaged cognitive control to the same degree as a more demanding cognitive task. In fact, there was some evidence of a decrease in relative alpha power from the first compared to second and third video viewing, though this was observed in the test session only.

3.4.2 Test-retest reliability

Test-retest reliability analyses revealed considerable variation in the reliability of EEG power depending on factors such as frequency band, power type (i.e. absolute or relative) and brain region. For average values across frontal and posterior brain regions, average theta power over the whole video showed fair to excellent reliability, whilst average alpha power over the whole video showed moderate and some poor reliability. Differences between brain regions averaged over the whole video instead showed the opposite pattern, with alpha measures tending to show fair (with some poor) reliability, whilst theta measures had largely poor reliability. Nearly all measures using the difference in average power between first and second video halves were not reliable, with only relative theta in some social conditions showing fair reliability. Intercepts from linear mixed models, taken as an additional measure of average power at video start, showed a broadly similar pattern to average power over the whole video, though notably the difference between frontal and posterior brain regions for both absolute and relative theta were poor. Reliability of residual standard error for all theta power measures was poor, whilst variability of absolute and relative alpha power averaged across the whole head, and of relative power in the frontal region were fair, indicating some consistency in the variability of alpha power.

Overall, relative measures were generally more reliable than absolute measures, and measures averaged across brain regions were more reliable than the difference between regions. Relative theta over the whole head and whole video were the most reliable measures, with mostly excellent reliability scores, though there is considerable variation in these. Moderate to good reliability values are in line with previous work (Haartsen et al., 2021; van der Velde et al., 2019) and as predicted here. This works provides information about the test-retest reliability of numerous measures, including sizeable disparities in the reliability of different measures. This information will be useful for EEG researchers when considering what measure to choose and how this may contribute to their results and interpretations.

3.4.3 Limitations

One potential limitation of the current study design is the low density of EEG channels covering the scalp, which was chosen to enable collection of good quality

EEG data which is notoriously difficult in this age range (as explained in section 3.2.2.1). Whilst low-density systems are also advantageous due to their portability, scalability, low cost and potential for use in everyday settings (Lau-Zhu et al., 2019b), high-density systems provide opportunities for analyses (i.e. connectivity) that are not possible here and may provide more reliable measures than the current system. Future work could benefit from investigating how reliability differs across EEG systems and with different densities of electrodes. Additionally, this study includes a sample of children for whom neurocognitive work is limited and utilises videos typically used with younger infants (Jones et al., 2015), meaning the current paradigm may not be most useful for work of this kind. Methods which optimise stimuli presentation for maximum neural response may lead to greater reliability by reducing the impact of inter-individual differences in brain development (Gui et al., 2022).

3.4.4 Conclusion

Research which investigates neural and cognitive development during the first few years of life is integral for improving scientific understanding about how development occurs and has practical implications for enhancing development where this is sub-optimal. Neural methods in particular hold great potential in improving current understanding of how early experiences impact development and may provide mechanistic insights about this relationship which cognitive and behavioural methods cannot. Despite such substantial potential, there remains progress to be made in conducting research which can be used as the base of such developmental support.

This study utilises a low-density, portable, and low-cost neuroimaging system which may improve accessibility to neurocognitive research and facilitate large-scale neuroimaging studies in low resource settings, which could be particularly impactful for participants including toddlers and low-SES families. Here, alpha and theta power measures during video viewing were specifically focussed on, as these are thought to be involved with cognition and learning (Begus & Bonawitz, 2020; Braithwaite et al., 2019; Klimesch, 2012; Orekhova et al., 2006), and are increasingly used in developmental work. Numerous measures of theta and alpha power were calculated, with significant differences in EEG power found between different frequency bands, brain regions, video conditions and other comparisons. Reliability for these measures varied considerably, with relative theta power over the whole head and whole video providing the most reliable measures. Researchers are therefore recommended to carefully consider reliability values when choosing measures in future studies.

4 TOOLS IN THE REAL WORLD: FEASIBILITY OF A FREE PLAY EEG DESIGN WITH TODDLERS

Abstract

Toddlerhood is a time of rapid development, with advancements at this age marking it as an important period to study. Despite this, neurocognitive research involving toddlers in currently somewhat lacking, likely due to practical difficulties with data collection association with toddlers' attention, motor, and cognitive abilities. This chapter seeks to help fill this gap by developing a paradigm designed specifically to gather good quantity and quality electroencephalography (EEG) data from toddlers. It uses a portable EEG system to record neuroimaging data during naturalistic paradigms including a free play session, meaning this design is additionally suited to use in less-controlled settings. 17 2- and 3-year-olds recruited from the Birkbeck Babylab and Toddlerlab database took part in an EEG study consisting of two parts; in part one participants watched live interactions with a researcher that included three conditions (bubble blowing, social and non-social), whilst part two involved a free-play session. During the free play session, children could explore four table-top activities, meaning this study was well-suited to investigate exploratory behaviours as well as evaluating the feasibility of a less-controlled design which may be well-suited for use with toddlers and in settings outside of the lab. Analyses revealed that theta power was higher during social and exploration, compared to non-social and bubble blowing conditions, whilst alpha was lower during exploration compared to all other conditions. These findings are interpreted in terms of literature implicating theta in active learning and alpha in inhibitory control. Relations were also found between depth (length) of exploration and each of theta and alpha in posterior regions, whilst a positive relation between experience of chaotic environments and depth of exploration may be considered in context of useful adaptations to experiences. Feasibility analyses showed high parental acceptability and suitability of this less-controlled design for gathering neurocognitive data from toddlers. This chapter sets out a blueprint for a study design which may be used in community settings in future research, thereby reducing some barriers to research for families and enabling more toddlers to participate in neuroimaging research.

Situate in thesis

Similar to the previous chapter, this chapter utilised a portable neuroimaging system which is suited for use in field or community settings. It assessed the feasibility of a less-controlled study design which was devised specifically to facilitate inclusion of toddlers and as a method which could easily be replicated in community settings. This design could additionally enable investigations about how children explore their environments and considers specific neural responses which may be particularly influenced by children's early experiences. The aim of this chapter was to assess the feasibility of using a portable neuroimaging system in a naturalistic experimental design, such that future studies can utilise a similar set-up in work which involves a diverse range of toddlers.

4.1 INTRODUCTION

Children undergo rapid motor, language and cognitive advances during toddlerhood, indicating this as a period of substantial development (Calkins, 2007; Colson & Dworkin, 1997). This progress is not only important during toddlerhood but also for later in life, with various evidence supporting a link between learning and abilities during toddlerhood with cognitive capabilities later in childhood (Fitzpatrick & Pagani, 2012; Rowe et al., 2012). In addition, the brain is particularly sensitive to environmental experiences early in life with evidence that differences related to socioeconomic status (SES) – a measure of someone's standing in society - are already apparent during the preschool age. Despite such motivation for research which focusses on neuro-cognitive development during toddlerhood, this remains relatively sparse, a fact likely due, at least in part, to practical difficulties with data collection with toddlers.

Whilst development during toddlerhood marks this as an interesting period to research, abilities during this age also contribute to challenges with gathering ample valid and reliable data. Toddlers' improving language abilities may not afford consistent and dependable responses, with potential for difficulties in both comprehension and production. Still-developing attention and self-regulation skills mean toddlers' focus on a single task is limited, whilst newly improved motor skills motivate physical exploration of their environment. Traditional experimental paradigms often require completion of standardised tasks and/ or attention to numerous repetitions of controlled stimuli and are particularly un-engaging for toddlers. Thus, one way to improve data collection from toddlers may be to utilise a more naturalistic design which enables children an element of free exploration, which could additionally improve ecological validity of the work.

4.1.1 Exploratory behaviours

Active exploration is a fundamental mechanism for how young children learn about the world around them, often outlined as motivated by a goal to gather information about a setting, object or person (Meyer, 1998; Pellegrini & Smith, 2001; Rusher et al., 1995; Weisler & McCall, 1976). Despite this clear goal, exploration may begin unintentionally with an initial spontaneous movement leading to learning, which causes motivation for subsequent intentional exploration (van Liempd et al., 2018). This may be particularly pertinent for toddlers whose newly developed motor skills facilitate extensive physical exploration of the world on their own terms. In addition to motor abilities, children's attention capacities may also be related to how they explore their environment, with strategies around allocation of attention and switching between activities forming crucial elements of exploratory styles (Blanco & Sloutsky, 2020).

Specifically, young children tend to distribute their attention relatively broadly, compared to adults who focus more selectively (Blair et al., 2009; Deng & Sloutsky, 2016). Though commonly considered the result of not-yet-developed executive attention abilities, more widely distributed attention may facilitate broader exploration, which could in turn lead to enhanced development. This is supported by evidence that children's early exploratory behaviour is related to later cognitive abilities, including spatial memory at 6 years (Oudgenoeg-Paz et al., 2014) and academic achievement at 10 years (Bornstein et al., 2013). As Blanco and Sloutsky (2020) suggested, patterns of attention allocation may be advantageous or disadvantageous depending on context. In turn, it is possible that differences in exploratory styles may be adaptive and based upon prior experiences. For example, there is evidence that children who had previous experience of institutional care showed less exploration (probing options with a range of potential outcomes) and greater exploitation (investigating limited but reward-promising options) compared to children who had not experienced institutionalisation, with this tendency towards more exploitation mediated by greater separation anxiety (Humphreys et al., 2015). It was argued that this evidence showed that anxiety may play a role in guiding an adaptive strategy which biases greater exploitation and less exploration, which may be advantageous in the context of early adversity. Given these findings, it may be fruitful to further investigate associations between early environmental experiences, anxiety, and exploration to understand whether similar relations occur in a sample with less extreme adversity.

In addition to children's own abilities, it is thus clear that opportunities for exploration and exploratory behaviours may depend on other factors including adult input and characteristics of the environment (Mayes et al., 1993; van Liempd et al., 2018). Studies investigating exploratory behaviours should thus carefully consider how to handle such factors, as well as planning how to measure exploration itself. Whilst much work has manipulated exploration as a particular condition within a task design, some work has used a more naturalistic measure of exploration. Such studies include those that have developed and utilised measures of exploratory behaviours such as the Exploratory Behaviour Scale (EBS) (Van Schijndel et al., 2010), whilst others assessed the depth and breadth of exploration during free play (van Liempd et al., 2018). In the latter study, depth related to the intensity of exploration and was measured by the amount of time a child was engaging in playing with an object, whilst breadth related to variety of exploration and was measured by the number of different uses of an object. Such differences in measures may lead to different findings relating to exploration; it is therefore important that results are interpreted carefully, and moves are made towards a study design which enables different measures to be explored. Nonetheless, existing work shows that exploration is, in itself, an important mechanism for young children's learning and development. Exploration also facilitates a more naturalistic study design which could lead to greater ecological validity of measures and improved data collection from toddlers.

4.1.2 Need for neuroimaging research

Whilst behavioural methods of exploration have revealed important information about children's development, neuroimaging may offer additional insights which cannot be observed behaviourally. This may be because neural development precedes behavioural development or could reveal a different neural mechanism underpinning analogous behaviours. Neuroimaging methods may be further advantageous as they could minimise potential confounders of cultural or societal differences which might impact behavioural measures (Lloyd-Fox et al., 2014).

Electroencephalography (EEG) is one such neuroimaging method which measures electrical oscillations in the brain, which are thought to be associated with information processing. EEG is a non-invasive method which can record brain activity during performance of tasks or activities with high temporal resolution, meaning it can provide information about rapidly altering patterns of brain activity underpinning behaviours or cognitive processing. Despite an established history of research using EEG with human participants, it is only relatively recently that wireless and portable systems have been developed. Wireless systems are detached from fixed appliances and may be particularly helpful in overcoming some practical difficulties involved with data collection from toddlers. Once applied to the head, participants can freely move around whilst wearing an EEG cap. Free movement enables young children to have breaks during an experiment without necessarily having to end EEG data collection as well as facilitating paradigms involving physical play and exploration. Both of these factors may lead to improved quantity and quality of data collection as they allow children to remain comfortable and engaged, in addition to enabling neural data collection during more naturalistic behaviours than traditional EEG systems.

A further potential benefit of portable systems is the opportunity to expand EEG research into community or field settings. This may be particularly impactful as a tool for improving diversity of neurocognitive research, as community-based research may reduce some barriers associated with participation in research for families of a particular background. Lab-based studies require participants to travel to often unfamiliar settings and demand resources which may be challenging for families of low-SES; as such, current research is typically limited by biased, unrepresentative samples. Portable systems which can be used in studies in community settings may thus facilitate participation of families from lower resource settings, enabling more inclusive and representative research. This is particularly important when considered in the context of work linking experience of lower resource settings with a higher risk of poor cognitive outcomes (as discussed in chapter 1). Increasing diversity of participants in neurocognitive research would lead to more generalisable findings and could be particularly enlightening in elucidating the relation between environmental experiences and development. Before portable EEG systems can be used in more naturalistic experimental designs in settings outside of the laboratory, feasibility must first be assessed; that is the focus of the current chapter. This study will add to existing work which has explored feasibility and acceptability of collecting EEG data in homes and field settings (Bhavnani et al., 2022; Lockwood Estrin et al., 2022; Troller-Renfree et al., 2021) by providing a study design suited to use with toddlers and in less-controlled settings. Where other researchers have found challenges in holding children's attention to videos as used in the lab when in a more interesting home environment (Troller-Renfree et al., 2021), a free play design may better engage participants.

In addition to assessing feasibility of design, there is motivation for investigating empirical findings relating to specific EEG measures. In particular, EEG power in range of frequency bands termed theta and alpha may be useful to explore. Theta and alpha power are both implicated in information processing, with work supporting a role of theta in memory, social attention and emotional processing (Guderian et al., 2009; Z. Jiang et al., 2017; Jones et al., 2015) and linking alpha to inhibition and visual attention (Foxe & Snyder, 2011; Klimesch, 2012; Klimesch et al., 2007). Given these links to executive attention, and the proposed link between attention and exploration, it follows that theta and alpha activity may also be related to specific exploratory behaviours.

In fact, some work already supports an association between theta oscillations and exploration or active learning. In adults, a positive relation between theta power and mean exploration time was found at the end, but not the start, of a period of exploration (Grunwald et al., 2001), whilst higher frontal theta power has been found during active compared to passive exploration (Chrastil et al., 2022). Further, work has suggested that mid-frontal theta oscillations may guide individual approaches to exploration, as relations were found between frontal theta and uncertainty, particularly in individuals who used uncertainty to guide exploration (Cavanagh et al., 2012). Given that this latter finding was made during participant-led free exploration, and that exploration might be considered as self-motivated information sampling, it has been hypothesised that theta oscillations may play a role in establishing optimal conditions for information processing (Begus & Bonawitz, 2020). In infants, frontal theta power during object exploration predicted later recognition of those objects (Begus et al., 2015), which Begus and Bonawitz (2020) suggested further supports a role of theta activity in establishing an individual's optimal learning condition. Based on this, they further suggested that theta activity might predict *what* children choose to explore, in addition to *how long* they explore for.

Other work has found frontal theta increases during exploratory behaviour in both infants and preschoolers (Orekhova et al., 2006). In this work, theta and alpha power was collected during three conditions: bubble blowing (baseline), speaking/ storytelling (social) and object manipulation (exploration). For both infants and preschoolers, higher theta activity was found during social and exploration conditions compared to baseline, though rate and regional differences were found. Specifically, greater increases were found in infants compared to preschoolers, with increased theta power predominantly frontal for infants but more widespread for preschoolers, possibly indicating a shift between these ages. In preschoolers, theta activity during exploration was predominantly over frontal areas but over parieto-occipital regions during attention to social stimulation. This work also found similar results in the alpha frequency band, with both infants and preschoolers showing lowest alpha power during the exploration condition, followed by the social and finally baseline condition. Given existing findings implicating alpha power in maintenance of focussed attention (Orekhova et al., 2001), and the hypothesised relation between attentional control and exploration (Blanco & Sloutsky, 2020), it follows that there may be a link between alpha power and exploration. Specifically, higher posterior alpha power in certain neural regions may be associated with greater engagement of inhibitory control (Klimesch et al., 2007), therefore it might be predicted that posterior alpha would be lower during periods of exploration, when children employ distributed attention strategies, rather than focussed attention.

Whilst EEG research has traditionally considered average power over a time period (i.e. 30 or 60 seconds) and compared this between conditions, there is growing research that considers changes at the moment-to-moment level. For example, there is evidence that analyses which use measures of change or variability may offer different but potentially useful information about neural processing (Garrett et al., 2013). Work in infants and young children, for instance, has found that increases in frontal theta power over the course of video viewing is related to cognitive ability at later ages (Braithwaite et al., 2019; Jones et al., 2020). Existing findings about theta change have involved infants aged 6 and 12 months and have not, to the best of knowledge, been extended to toddlers. Given the limitations of findings about theta change, work indicating that activity in the theta and alpha frequencies may be related to exploration, and suggestions that moment-to-moment changes in power may be informative (Garrett et al., 2013), changes in EEG power over the course of a toy engagement and relations between change and exploratory behaviours were also explored in the current work.

4.1.3 The current study

In the current study EEG was collected during conditions designed to be more naturalistic than typical screen-based studies. Part one of the EEG session involved live interactions with a researcher including three conditions: bubble blowing, social and non-social interactions. These conditions were based upon those commonly used in developmental neurocognitive research and adapted from traditionally-used screen-based videos (i.e. Jones et al., 2020), designed to be more naturalistic in order to improve ecological-validity of findings. The second part of the EEG session utilised a free-play protocol during which children could explore four table-top activities, designed to investigate free exploration as well as evaluating the feasibility of a less-controlled design which may be well-suited for use with toddlers and in settings outside of the lab. Table-top activities were chosen to enable children exploration opportunities whilst reducing the impact of motor artifacts on EEG data; boards for the activities were attached to the table meaning children's movements were somewhat limited during engagement with an activity. It was expected that this design would enable a lower attrition rate of EEG data collection than traditional, more-controlled studies and would provide a feasible protocol which can be used with toddlers in a range of settings. Feasibility was assessed by considerations of the quality and quantity of data collected, as well as tolerance measures and parent-reported feedback about the study, collected via a parent feedback questionnaire.

In line with previous work, predictions were also made about theta and alpha power during the three conditions: bubble blowing, social, non-social and toy engagement. It was hypothesised that theta power would be higher in the social and toy engagement conditions compared to both non-social and bubble blowing conditions, similar to Orekhova et al. (2006). As children in the current study were between the ages of the two groups in that study, which found regional differences between age groups, firm regional predictions were not made but were explored. By contrast, it was predicted that alpha would be lower during toy engagement compared to all other conditions. This is based on work indicating a role of alpha in inhibitory control (Klimesch, 2012), and given that distributed attention may be important for exploratory behaviours (Blanco & Sloutsky, 2020), whereas more focussed attention is required during the bubble blowing, social and non-social conditions. Based upon these findings and supported by work which found that theta power preceding visual fixations predicted duration of subsequent fixation in 12-month-olds (Wass et al., 2018), it was further hypothesised that there may be a positive association between theta power and exploration, and a negative association between alpha power and exploration during free play. Given that some existing findings have found differences in theta power over multiple repetitions of a stimulus (Meyer et al., 2019), condition analyses were first conducted using each child's first toy engagement only, and then repeated using all bouts of toy engagement. This enabled investigation into whether the pattern of neural activity differed between first and all engagements, whilst including all engagements additionally allowed for greater statistical power. The main measure of exploration for these hypotheses was time of engagement, which is in line with the depth measure of exploration used by van Liempd et al. (2018). Other measures of exploration, such as grouping by exploratory style, could be extracted from this study design in future work but were limited by sample size in the current study.

The association between exploration and measures of anxiety and early environments was also explored to understand whether higher anxiety and greater experience of adversity was related to reduced exploration as in Humphreys et al. (2015). The General Anxiety Disorder (GAD-7) and behavioural inhibition system (BIS) questionnaires were used to measure anxiety levels in parents and children respectively; both questionnaires have been found to have good psychometric properties (Carver & White, 1994; Leone et al., 2001; Ruiz et al., 2011; Spitzer et al., 2006). Given that unpredictability and threat have both been cited as important environmental characteristics which may relate to experience of adversity (Ellis et al., 2022), the Confusion, Hubbub and Order Scale (Home CHAOS), which measures confusion and unpredictability in a household (Matheny et al., 1995), and neighbourhood safety scale, as an index of safety/ threat in a community (Mujahid et al., 2007), were used.

4.2 METHODS

4.2.1 Participants

Participants were 17 (7 female) toddlers recruited from the Birkbeck Babylab and Toddlerlab database that parents had previously signed up to. Parents were contacted when children were aged between 24- to 48-months-old and invited to take part in the study (mean age at visit = 33.9 months, *SD* = 5.4). Information was sent via email and a mutually suitable time was arranged for a visit to the Birkbeck Toddlerlab. Consent for this study was gained from Birkbeck, University of London; ethics number 2122044. Participants received a ‘young scientist’ t-shirt after taking part in the study.

Demographic information for participants included in this study are in Table 4.1 and Table 4.2 below. One parent did not complete the questionnaires due to challenges during their visit; they were asked to complete the questionnaires online after their session but chose not to. Eight (50.0%) of the remaining parents had studied at postgraduate level, six (37.5%) had studied at undergraduate level and the remaining two (12.5%) had completed further education (i.e. A levels, BTECs or equivalent). Seven parents were currently working full time, five were working part time, two were currently looking for work, and two parents chose not to answer. Attending parents were also asked if they were comfortable answering some questions about their child’s other parent; all eleven of those who responded said that the other parent was working full time.

Table 4.1: Demographic data for the recruited sample

	Mean	SD	Min.	Max.	*N*
Participant age at visit (months)	33.9	5.4	26	44	17
Parent 1 years in education	18.1	2.2	13	22	16
Parent 2 years in education	17.9	2.8	13	22	12
Ladder: standing in UK	6.2	1.6	3	8	16
Ladder: standing in community	6.1	1.2	4	8	16
Number of bedrooms in household	2.6	0.8	1	4	15
Number of people in household	2.9	0.7	2	4	16

Table 4.2: Descriptive statistics for household income for the recruited sample

	Category	*N*	Percent
	Less than £20,000	1	6.7%
	£20,000 - £29,999	0	0%
	£30,000 - £39,999	0	0%
Annual household income	£40,000 - £59,999	2	12.5%
	£60,000 - £79,999	1	6.7%
	£80,000 - £99,999	2	13.3%
	£100,000 - £149,999	0	0%
	More than £149,999	6	40.0%
	Do not wish to answer	3	18.8%

4.2.1.1 Data inclusions

Of the 17 recruited participants, three participants provided no usable data; of these three, two refused to wear the EEG cap, and Matlab issues for one participant meant the recorded EEG data was unusable (see Figure 4.1). Some data from the remaining 14 participants was not usable due to equipment failures in a particular condition, but all 14 participants contributed data to at least one condition (see Table 4.3 for numbers per condition). Social and non-social interactions were repeated three times each; some participants only provided data for one or two repetitions due to technical issues. One participant took part in the free play session but did not engage with any toys on the tabletop; this participant was therefore excluded from free play analyses. All usable data provided by participants were included in analyses as far as possible.

Figure 4.1: Exclusion information for the recruited sample

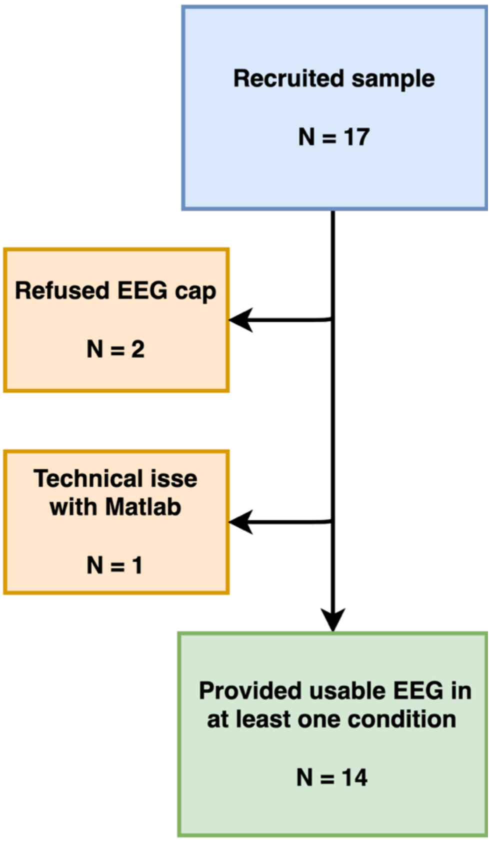


Table 4.3: Number of participants providing some usable EEG data for each condition

	N
Whole sample	17
Bubble blowing	13
Social interaction	13
Non-social interaction	13
Initial engagement in free play	13
Multiple engagements in free play	14

4.2.2 Materials and stimuli

4.2.2.1 EEG session

4.2.2.1.1 Live interactions

During live interactions a researcher was seated at a table opposite the participant who was either seated alone or on their parent's knee. Bubble blowing was accomplished by the researcher using standard soapy bubbles. The social interaction condition was characterised by nursery rhymes performed by the researcher (see Table 4.4 for details), whilst the non-social condition consisted of mechanical toys being set off by the researcher. The order of toys and rhymes within each 1-minute repetition was not fixed.

Table 4.4: Nursery rhymes and gestures performed during social interactions

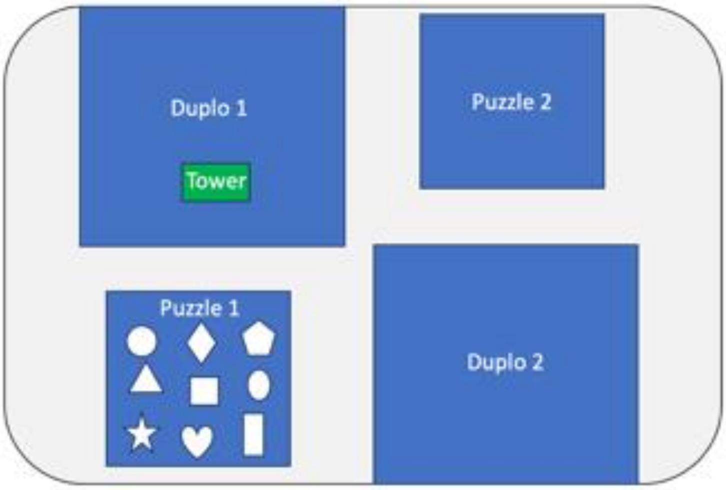
	Lyrics	Gestures
"Hi Baby"	Hi Baby!	
	Where are my eyes?	Question gesture
	Here are my eyes	Point to eyes
	Where is my nose?	Question gesture
	Here is my nose	Point to nose

	Where is my mouth?	Question gesture
	Here is my mouth	Point to mouth
Wheels On the Bus	The wheels on the bus go round and round	Rotate hands round and round….
	Round and round	Rotate hands round and round….
	Round and round	Rotate hands round and round….
	The wheels on the bus go round and round	Rotate hands round and round….
	All through the town	Wave both hands side to side
Incy-Wincy Spider	Incy-Wincy Spider went up the waterspout	Pincer grip, spider climbing up
	Down came the rain and washed the spider out	Wriggle fingers for rain coming down
	Out came the sunshine and dried up all the rain	Move both hands to mimic the sun, then gesture upwards
	So Incy-Wincy spider went up the spout again	Pincer grip, spider climbing up
Twinkle-Twinkle	How I wonder what you are?	Question gesture
	Up above the world so high	Point towards the sky
	Like a diamond in the sky	Make a diamond shape with hands
Pat-A-Cake	Pat-a-cake pat-a-cake baker’s man	Clap hands from side to side
	Bake me a cake as fast as you can	Gesture mixing a cake in a bowl
	Roll it	Gesture rolling a mixture
	And pat it	Clap hands from side to side
	And mark it with a B	Draw a “B” with index finger
	And put it in the oven for	Pretend to place cake in an oven
	Baby and me!	Point to child then point to self

4.2.2.1.2 Free play session

The room was set up as in Figure 4.2. A child-level table was in the centre of the room, with four activity stations on top; two were DUPLO® activities and two were puzzle activities. The DUPLO® stations each had a 15inch x 10.5inch base board with 16 loose bricks (2 squares and 2 rectangles of each of the following colours: blue, red, yellow, green) laid atop. The only difference across the two stations was that one had a DUPLO® tower already build alongside the blocks (DUPLO® 1), whilst the other had nothing (DUPLO® 2). The puzzle stations each consisted of a 20cm x 20cm wooden board with 9 coloured shapes, which were identical across the two puzzles. The only differences were that one puzzle had indented spaces for the coloured shapes to fit (puzzle 1), whilst the other was a flat board (puzzle 2). These tasks were designed to have one creative and one aim-based station for each activity type. The configuration of the table could be either as in Figure 4.2a, where puzzle 2 was closest to where the child entered the room or in reverse (as in Figure 4.2b) where puzzle 1 was closest to entry.

Figure 4.2: a) Lay-out of tabletop from above and b) picture showing table with activity stations in the laboratory

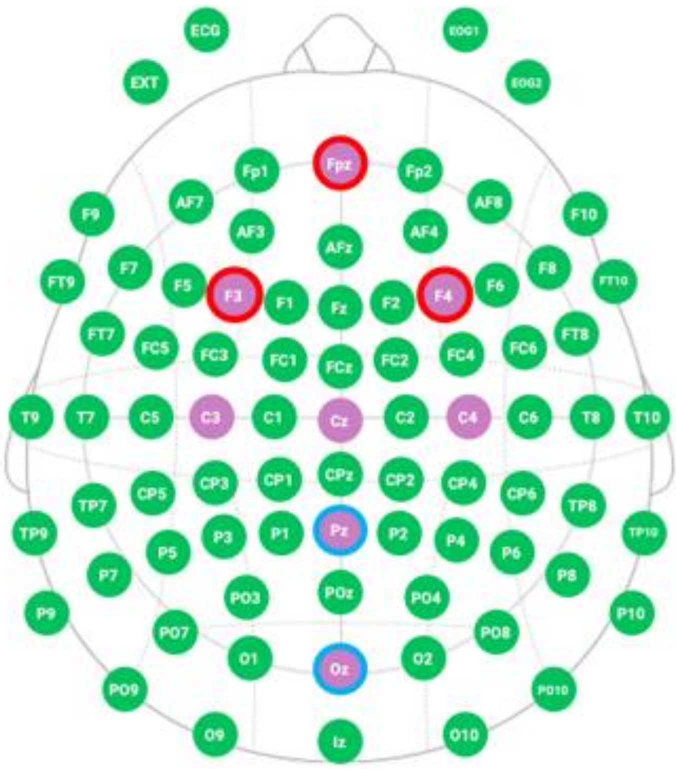


4.2.2.1.3 EEG recording

EEG was recorded using a wireless geltrode Enobio EEG system (NE, Neuroelectrics, Barcelona, Spain). Data was recorded from eight electrodes placed at FPz, F3, F4, Cz, C3, C4, Pz and Oz (see Figure 4.3 for channel layout) and transferred to a MacBook Pro via a Wi-Fi connection. A low-density array was used to facilitate good quality data collection by balancing this with quantity of channels. Young children typically have a considerable amount of hair, meaning some time is required to seat electrodes and ensure data quality, however they also have limited patience and often low tolerance during application, meaning application time needs to be kept to a minimum. Using large arrays can lead to greater attrition rates, as some children's patience may be surpassed and they may end up providing no data, whereas this could be achieved when using a smaller array. These channels were chosen such that a frontal and a posterior region could be derived, whilst Cz, C3 and C4 were used for re-referencing (4.2.4.1) in line with chapter 3.

The CMS and DRL reference electrodes were attached to an ear clip placed on one of the participant's ears. Before applying the ear clip, the area was gently cleaned with antibacterial spray. The software Neuroelectrics NIC 2.0 (Barcelona, Spain) was used to record data with a sampling rate of 500Hz and to visually monitor data quality throughout the session. The Neuroelectrics Quality Index (QI) calculates line noise (power in the range of 49 - 51Hz), main noise (power in the range of 1-40Hz) and the offset of the signal every 2 seconds.

Figure 4.3: Layout for Enobio cap; channels used in this study are coloured purple



4.2.2.2 Questionnaires

Parents competed a total of seven questionnaires in this study, with only data from the Home CHAOS (Matheny et al., 1995), the neighbourhood safety scale (Mujahid et al., 2007), the GAD-7 (Spitzer et al., 2006), the BIS (Carver & White, 1994) and a parental feedback questionnaires included in the current study.

The Home CHAOS is a 15-item forced-choice questionnaire designed to measure home confusion and disorganisation (Appendix C.1). For each statement, participants responded on a scale of 1 (very much) to 4 (not at all). The neighbourhood safety score contains 16 items detailing statements relating to crime and safety, and physical and social elements of their neighbourhood (Appendix C.2). Each item was answered on a 10-point Likert scale ranging from 1 (rarely/ not worried) to 10 (frequently/ very worried).

The GAD-7 is a measure of parental anxiety, in which parents answered seven items according to a 4-point scale scored from 0 to 3 (Appendix C.3). The BIS is a measure of child anxiety and contained seven items relating to worries and fears (Appendix C.4). Parents responded to each item according to a 4-point scale scored from 1-4. A parent feedback questionnaire was also designed to assess parental judgement about the feasibility and suitability of the experiment for their child. The current study considers ratings about the study design which parents provided based on a 5-point scale (Appendix C.5).

4.2.3 Procedure

After consent was gained from a parent/ guardian, the child was fitted with an EEG cap and the EEG session began. There were two main parts to the EEG session: (1) live interactions and (2) a free play session. Following this, the EEG cap was removed, and children completed a cognitive session involving three tablet tasks followed by two play-based cognitive assessments, though these measures are not included in the current chapter.

4.2.3.1 EEG session

During EEG cap application, participants were shown child friendly cartoons (such as Peppa Pig) on a tablet and entertained by one researcher, whilst another applied the

EEG cap. The EEG cap had previously been prepared so that it contained geltrodes in each of the eight electrode locations used in this study. The cap (with geltrodes inserted) was placed onto the participant's head and the chin strap fastened. The researcher then inserted gel to each of the eight geltrodes, before attaching electrodes to each. The earlobe was cleaned and the earclip (to which reference electrodes were attached) was applied. The electrode bundle was then plugged in to the NIC box and the programme NIC2.0 was used to visually assess the EEG signal. Once the EEG signal was deemed good enough, live interactions began. Following cap application, a measure of capping tolerance was noted on the participant's session sheet. Tolerance was rated on a 5-point scale, with a score of 1 indicating complete acceptance and a score of 5 reflecting total refusal.

4.2.3.1.1 Live interactions

Children were seated across a table from a researcher and asked to watch carefully. Parents were seated at the side of the room, just next to their child, unless the child got upset in which case the parent moved and the child sat on their knee. A researcher blew bubbles for one minute (bubble blowing condition), then started a live paradigm including a social and non-social condition. In the social condition the researcher sung nursery rhymes and in the non-social condition they played with mechanical moving toys. Each condition lasted for 1 minute and was repeated three times, totalling 6 minutes. Conditions were counterbalanced, with the order indicated by a Matlab script. Timings were also indicated by a Matlab script, which played a beep sound to indicate the start and stop of each bout to the researcher.

Whilst singing nursery rhymes in the social condition, the researcher maintained eye contact with the child and smiled throughout the task. They used hand gestures to accompany rhymes above the table at approximately chest height. When playing with toys in the non-social condition, the researcher was instructed to avert eye contact as they placed one toy a time on the table in front of the child and gave each toy 3 or 4 spins or presses before moving on to the next toy. There were 4 different toys and when not in use, toys were placed on a second table to the side of the researcher.

4.2.3.1.2 Free play session

Following live interactions, participants entered a second room and were told 'in the next room there are lots of toys on the table for you to play with. When we go in, you can play with any of the toys on the table.' Parents/ carers were asked to sit on the sofa at one edge of the room and were told 'please give your child as little input as possible whilst they are playing in this room, as we are interested in how they explore themselves. Obviously your child might look to you for reassurance – where possible please respond non-verbally (i.e. nodding, smiling) rather than saying anything. Of course, please do reassure them if needed and we might ask you to encourage them at a certain point, but this will be prompted. We have some questionnaires for you to complete and it would be great if you could do these while your child is playing.' Parents/ carers were given questionnaires for them to complete during this part of the study. One researcher knelt to the side of the table.

Children were allowed to play with any of the items on the tabletop. When children first engaged with an item, a key was pressed by the second researcher to indicate that this had happened. A child was considered to engage with an item when they actively touched or picked it up (i.e. accidentally brushing an item was not considered an engagement but reaching to touch an item was). If children were still playing with the same item after 2 minutes, a researcher asked if they would like to look at any other toys on the table. If a child switched to a different activity, they were left to play with it for two minutes, before being prompted to switch again. If a child did not switch tasks after the initial prompt, the researcher approached the table and indicated next to each activity 'in this activity you can…' and demonstrated with up to 3 pieces/ bricks. Whenever a child switched to a different task, a key was pressed to indicate this new engagement.

4.2.3.2 Questionnaires

Whilst children were taking part in tasks, their parent/ carer was asked to complete a series of questionnaires. Parents remained in the room with their child and researchers were on hand to explain questionnaires or to answer any questions they had. Three parents completed questionnaires on paper, whilst all other parents completed questionnaires electronically using Psytoolkit (Stoet, 2010, 2017) on a tablet device. The researcher entered a pseudo anonymised participant ID number before handing the device to parents to complete the questionnaires. Parents completed a total of seven questionnaires, which took around 25 minutes.

4.2.4 Data processing

4.2.4.1 EEG

4.2.4.1.1 Pre-processing

Data recorded during experimental conditions were extracted from the rest of the recording for pre-processing. Data were pre-processed using a combination of in-house written scripts and Fieldtrip Matlab scripts (Oostenveld et al., 2011) in Matlab R2021a. All datasets were checked and organised according to the available data for each participant. Enobio data were converted into Fieldtrip format for further pre-processing.

Data were split into separate files based on condition; for each file continuous data were then segmented into 1-second epochs with no overlap. Data were detrended and a 0.1-48 Hz bandpass filter was applied to filter out 50Hz line noise and higher-frequency noise from muscle artifacts. Artifacts were identified using both automatic and manual detection. In automatic detection, trials were marked unusable if the signal exceeded thresholds -150 to 150 µV, or if a flat signal or a jump of greater than 20 µV were detected; these trials were then removed. In manual detection, trials were removed if more than one channel displayed artifacts or if there were artifacts in the signal from any one channel in the particular regions of interest (i.e. Fpz, F3, F4, Pz or Oz).

Artifact-free data were then detrended and re-referenced on a trial-by-trial basis to Cz, or the average of C3 and C4 if Cz contained artifacts. If C3 or C4 also contained artifacts, the whole trial was excluded from further analysis. A fast-fourier transformation was applied to re-referenced data to ascertain power (μv^2) per electrode in 1Hz bins for 1-48Hz. Power data were split into conditions, and absolute and relative power were calculated. Relative power was calculated, as this had been found to be more reliable than absolute power (chapter 3). Relative power was calculated by dividing the sum of power in the theta (3-6Hz) and alpha (6-10Hz) frequency bands by the sum of power in the 2-20Hz range, as in Segalowitz et al. (2010).

4.2.4.1.2 EEG measures

A number of measures of power were used, with theta and alpha relative power measures calculated for each of the frontal and posterior regions. Measures in the frontal regions were taken as the average power over channels Fpz, F3 and F4 (circled in red in Figure 4.3), whilst the posterior region calculated power over Pz and Oz (circled in blue in Figure 4.3).

Overall average power was calculated as mean power over the duration of a condition, for any 1-second segments that were left after processing and artifact removal. For bubble blowing, social and non-social interactions, average power was calculated over the 60 seconds of each trial; for toy engagement during free play power was calculated over 30 seconds. This included the five seconds preceding engagement and twenty-five seconds post-engagement; this time period was chosen to ensure neural changes that occurred immediately before behaviour changes were captured whilst reducing the impact of motor artifacts which may have occurred earlier. A duration of 30 seconds was used to maximise inclusion of toy engagements as children often disengaged after this; 60 seconds were used for other conditions in line with previous work (i.e. Jones et al., 2020).

A more continuous measure of power over the course of video viewing was also calculated, using average power in each 1-second segment of video viewing to consider power change over segment number, in addition to other condition comparisons. Linear mixed models were fitted using power as dependent variable and segment number as a fixed effect. An intercept value was also obtained from this model for each subject, which were then used as a second measure of individual average theta power.

4.2.4.2 Exploration measures

Depth of exploration was calculated as length of engagement (in seconds) with an activity. These were calculated using EEG markers indicated by researchers during the test session and behavioural coding of videos after.

4.2.4.3 Questionnaire measures

A Home-CHAOS score was calculated as the total of scores on individual items, with a maximum of 60. A similar method was used for neighbourhood score, enabling a maximum score of 160. For BIS score the maximum score was 28, whilst a single GAD-7 score was calculated as the summation of scores on individual items enabling a maximum score of 21. Individual questions on the parental feedback questionnaire were scored from 1 to 5, whereby 5 indicated high and 1 indicated low satisfaction. Scores for individual questions were used in this chapter.

4.2.5 Statistical analyses

Condition effects were investigated using linear models fitted with average power as the dependent variable and independent variables of region (frontal or posterior), and condition (bubble blowing, social, non-social and toy engagement). Statistics from ANOVAs run on the linear models were reported as they enable more precise and powerful analyses compared to t-tests provided by the linear models; follow-up tests were pairwise comparisons. Analogous analyses were conducted for first toy engagements only and all occurrences of toy engagements; as a reminder, this enabled exploration into whether the pattern of neural activity differed between first and all engagements, whilst including all engagements additionally allowed for greater statistical power.

To extract measures of change in power over the course of a video viewing, a linear mixed model was used with power per 1 second segment as a dependent variable and segment number as a fixed effect. Participant ID was included as a random effect allowing for variable slope and intercepts; engagement number was also included as a random effect partially crossed with participant ID. Separate models were run for frontal and posterior theta and alpha, with random effects extracted per engagement for each participant from each model.

To investigate relations between neural measures and exploration, linear mixed models were fitted with EEG power as the dependent variable and a fixed effect of engagement length. Participant ID was included as a random effect allowing for variable slope and intercepts with engagement number also included as a random effect partially crossed with participant ID. Analyses were repeated using both average power and extracted random effects as a measure of power change.

Any trials which had fewer than 10 seconds of useable EEG were excluded from analyses. Assumptions were checked before models were fitted and other tests were used where there were concerns (these are outlined where appropriate in the results). Most statistical analyses were conducted in RStudio (Core Team, 2023), though some simple statistics were conducted using Matlab R2021a or SPSS 29.0.0.0; figures were all produced in RStudio.

4.3 RESULTS

4.3.1 Feasibility of less-controlled design

Descriptive statistics were calculated for the number of usable EEG segments in each condition for the 14 children from whom EEG data was collected (Table 4.5). Of note, measures for bubble blowing and live interactions are taken over 60 seconds, whilst free play measures are taken over 30 seconds (see section 4.2.5 for explanation). An issue with the reference electrodes meant one participant was excluded from bubble blowing, social and non-social interaction conditions, though this was rectified for the free play session. For one other participant, technical issues meant no end marker was recorded for bubble blowing, therefore they are excluded from these analyses. All other children provided a good amount of usable EEG data in all conditions. Mean number of usable segments of data is lower for the third repetitions of both social and non-social interaction conditions; this is largely due to technical issues which meant that data was unusable for three participants in the social and two participants in the non-social condition. When these participants were excluded, descriptive statistics for usable segments in the third repetitions were as follows: social; M = 36.5, SD = 16.5, lower range = 10, upper range = 59, N =10 and non-social; M = 44.2, SD = 13.5, lower range = 22, upper range = 60, N =11.

Researcher-rated tolerance of capping up suggest reasonable compliance (where 1 indicated high acceptance and 5 was total refusal) from children in this study (Table 4.5). Two participants who refused the EEG cap were scored as five, whilst 9 participants were scored as one or two, indicating minimal upset. All children who wore the EEG cap completed the EEG session, although one child did not interact with any toys during the free play session. All children wore the EEG cap for at least 20 minutes, with 13 children doing free play for at least 10 minutes.

Table 4.5: Descriptive statistics for usable segments of EEG per segment, tolerance of capping up and time spent wearing the EEG cap

		M	*SD*	LB	UB	*N*
Usable segments from bubble blowing		41.1	14.0	10	57	12
Usable segments from social	First repetition	41.4	15.4	16	59	13
	Second repetition	45.6	15.3	6	60	13
	Third repetition	28.1	21.5	0	59	13
Usable segments from non-social	First repetition	45.9	14.4	23	60	13
	Second repetition	45.8	16.3	10	58	13
	Third repetition	37.4	20.7	0	60	13
Usable segments from free play engagements		21.7	4.9	1	30	66
Usable segments from free play initial engagement		21.1	5.4	10	27	8
Researcher-rater tolerance of capping up		2.5	1.4	1	5	17
Time spent wearing the EEG cap		29m 58s	6m 33s	21m 47s	39m 46s	14

Parent feedback generally indicated good satisfaction with the study, with ratings dropping no lower than 4 out of 5 (where 5 was the highest possible satisfaction rating) on questions about study location, design, and importance: see Table 4.6. Parents reported that children's well-being, health and mood on the day of their participation was mostly representative of their typical state (*M* = 4.2, *SD* = 1.0, LB = 2, UB = 5, *N* = 13) and children were generally familiar with the toys used during the free play session, though were less likely to often play with an electronic tablet or similar (Table 4.7).

Table 4.6: Summary statistics for parent ratings of this study from a 1-5 scale

	M	*SD*	LB	UB	*N*
Location of the visit	4.8	0.4	4	5	16
Duration of the visit	4.7	0.5	4	5	16
Assessment of your child	4.9	0.3	4	5	16
The way the study and reasons for the study were explained	4.9	0.3	4	5	16
Importance of study according to the outlined aim	4.8	0.5	4	5	16
Rating of assessment as a way of assessing children's neurodevelopment from a parent's point of view	4.6	0.5	4	5	16

Table 4.7: Descriptive statistics for how often participants typically play with toys used in the study based on a 1-5 scale

	M	*SD*	LB	UB	*N*
Electronic tablet	2.9	1.3	1	5	16
Lego/ Duplo	4.4	0.6	3	5	16
Puzzles/ jigsaws	4.1	0.7	3	5	16

4.3.2 Condition effects of average EEG power

This section contains descriptive statistics and condition analyses for average alpha and theta power during live bubble blowing, toy playing, singing and initial engagement in the free play session. To maximise upon the available data, linear regressions were used to compare between conditions, as these do not require each participant to have data for every condition. Since some participants did not have EEG for all three repetitions of social and non-social conditions and as significant effects of video number were not found in the previous sample in chapter 3, analyses here are collapsed across repetitions. Given previous findings of a significant difference in alpha compared to theta power (see chapter 3), separate analyses were conducted for alpha and theta power.

Repeated-measures linear models were fitted using the *lm* function in RStudio (Core Team, 2023) with average power as the dependent variable and independent variables of region (frontal or posterior) and video condition (social versus non-social). Videos which had fewer than ten usable segments of EEG data were excluded from analyses and assumptions were checked. Where significant interactions were found, follow-up pairwise contrasts were conducted using the *emmeans* function in RStudio (Lenth, 2023) with a Bonferonni correction for p-values.

4.3.2.1 First engagements only

4.3.2.1.1 Alpha

An ANOVA of the linear model showed no significant effects at the $p < .05$ level (Table 4.8). Figure 4.4 provides a visualisation of group means, whilst Table 4.9 shows descriptive statistics for all conditions.

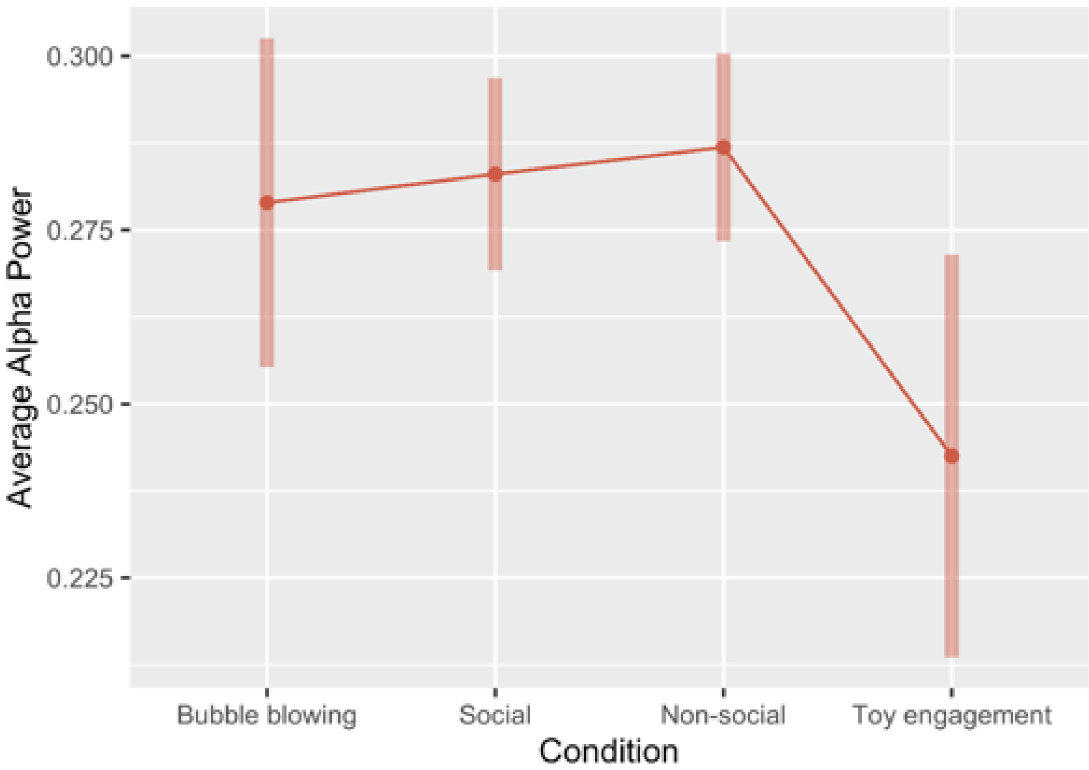
Table 4.8: F-value, p-value and partial eta squared effect size for ANOVAs on average alpha power across regions and conditions when only first engagements with toys were included

	F	p	η^2 (partial)
Region	0.49	.483	<0.01
Condition	2.58	.055	0.04
Region x Condition	0.62	.604	0.01

Table 4.9: Mean, standard deviation and N for average alpha power in the frontal and posterior regions for each condition when only first engagements with toys were included

Region	Condition	*Mean*	*SD*	*N*
Frontal	Bubbles	0.28	0.07	12
	Social	0.27	0.05	35
	Non-social	0.29	0.05	37
	Toy engagement	0.25	0.03	8
Posterior	Bubbles	0.28	0.06	12
	Social	0.29	0.07	35
	Non-social	0.29	0.06	37
	Toy engagement	0.24	0.05	8

Figure 4.4: Mean and confidence intervals for average alpha power over conditions when only first engagements with toys were included



4.3.2.1.2 Theta

An ANOVA of the linear regression found a significant main effect of condition (Table 4.10). Follow-up pairwise comparisons with Bonferonni correction found that theta power was lower during bubble blowing compared to each of the social interaction and toy engagement conditions (Table 4.11 and Table 4.12). Theta power during social interaction and toy engagement was also significantly higher than during the non-social condition (Table 4.11 and Table 4.12). Figure 4.5 displays these differences; Table 4.13 shows descriptive statistics for all conditions.

Table 4.10: F-value, p-value and partial eta squared effect size for ANOVAs on average theta power across regions and conditions when only first engagements with toys were included (star indicates significance)

	F	*p*	η^2 (partial)
Region	0.81	.370	<0.01
Condition	7.59	< .001*	0.11
Region x Condition	0.16	.927	<0.01

Table 4.11: Estimate, t-value, p-value, and standard errors for pairwise comparisons between average theta power during each condition when only first engagements with toys were included (star indicates significance)

		β	*t*	*p*	*SE*
Bubbles	Social	-0.05	-4.07	.0004*	0.01
Bubbles	Non-social	-0.02	-1.67	.343	0.01
Bubbles	Toy engagement	-0.05	-2.86	.024*	0.02
Social	Non-social	-0.03	-3.43	.004*	0.01
Social	Toy engagement	<-0.01	-0.14	.999	0.01
Non-social	Toy engagement	0.03	1.93	.219	0.01

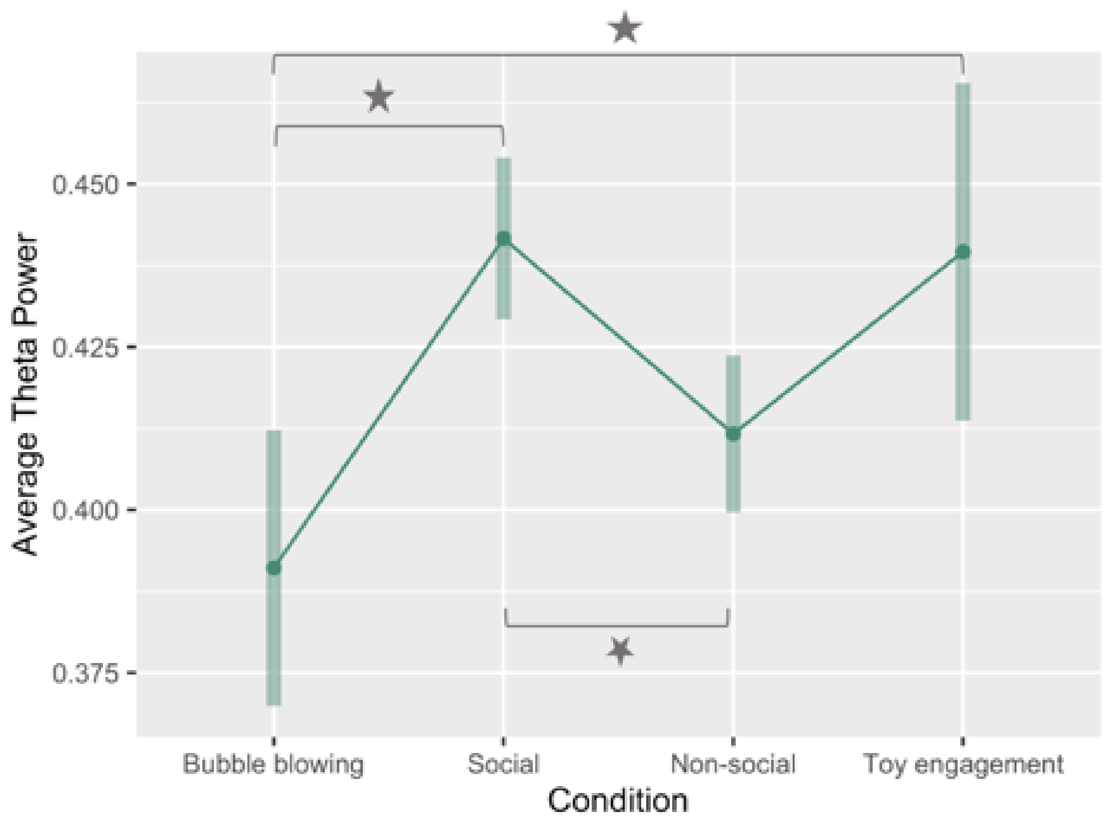
Table 4.12: Mean, standard error, and 95% confidence intervals for average theta power in each condition when only first engagements with toys were included

Condition	*M*	*SE*	LB	UB
Bubbles	0.39	0.01	0.37	0.41
Social	0.44	0.01	0.43	0.45
Non-social	0.41	0.01	0.40	0.42
Toy engagement	0.44	0.01	0.41	0.47

Table 4.13: Mean, standard deviation and N for average theta power in frontal and posterior regions across conditions when only first engagements with toys were included

Region	Condition	*M*	*SD*	*N*
Frontal	Bubbles	0.28	0.07	12
	Social	0.27	0.05	35
	Non-social	0.29	0.05	37
	Toy engagement	0.43	0.05	8
Posterior	Bubbles	0.28	0.06	12
	Social	0.29	0.07	35
	Non-social	0.29	0.06	37
	Toy engagement	0.45	0.07	8

Figure 4.5: Mean and confidence intervals for theta power over conditions when only first engagements with toys were included



4.3.2.2 All engagements

4.3.2.2.1 Alpha

An ANOVA of the linear model found a significant main effect of condition on alpha power (Table 4.14). Follow-up pairwise comparisons with Bonferonni correction found significant differences in alpha power in the toy engagement compared to each of the bubble blowing, social and non-social conditions (Table 4.15), with means indicating this was lower during toy engagement (Table 4.16). Figure 4.6 illustrates these differences, whilst Table 4.17 shows descriptive statistics for all conditions.

Table 4.14: F-value, p-value and partial eta squared effect size for ANOVAs on average alpha power over regions and conditions including all toy engagements (star indicates significance)

	F	*p*	η^2 (partial)
Region	0.93	.336	<0.01
Condition	15.67	< .001*	0.13
Region x Condition	0.64	.590	0.01

Table 4.15: Estimate, t-value, p-value, and standard errors for pairwise comparisons of average alpha power (including all toy engagements) across conditions (star indicates significance)

		β	*t*	*p*	*SE*
Bubbles	Social	-0.004	-0.33	0.987	0.01
Bubbles	Non-social	-0.01	-0.65	0.914	0.01
Bubbles	Toy engagement	0.04	3.13	0.010*	0.01
Social	Non-social	0.004	0.45	0.970	0.01
Social	Toy engagement	-0.04	-5.23	<.001*	0.01
Non-social	Toy engagement	-0.04	-5.3	<.001*	0.01

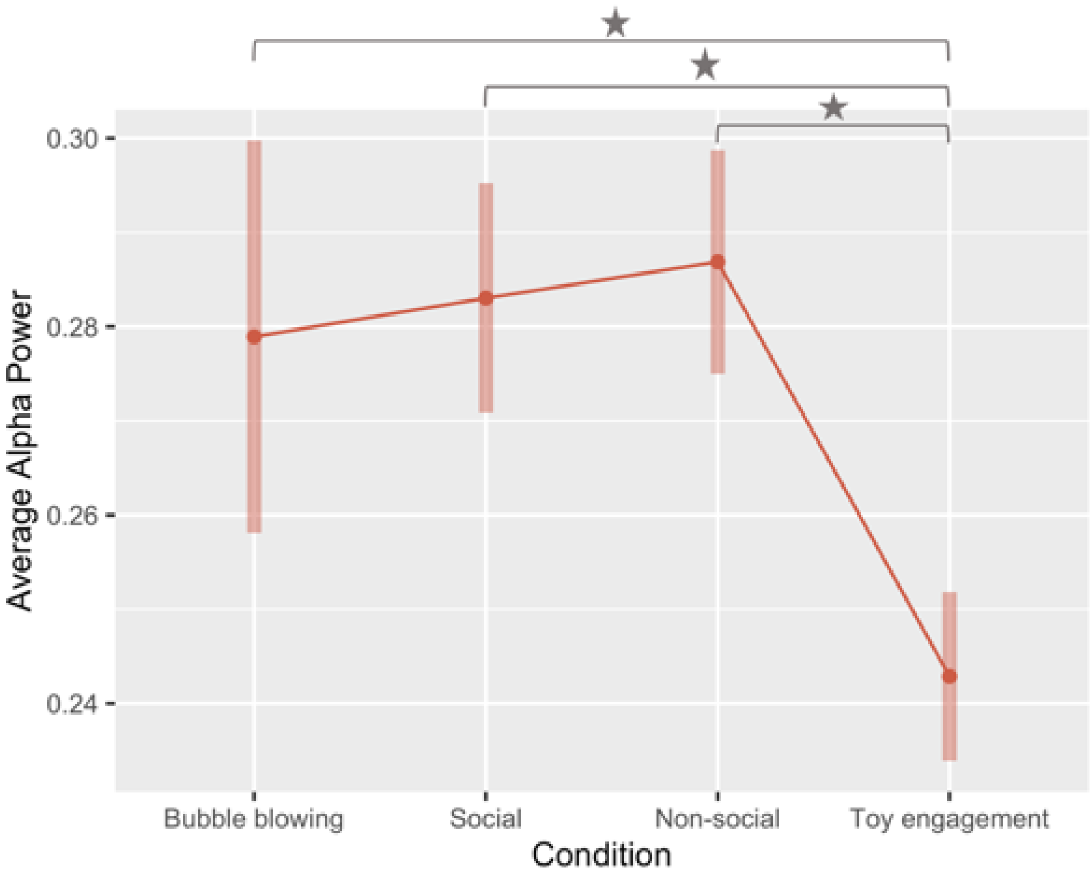
Table 4.16: Mean, standard error, and 95% confidence intervals for average alpha power including all toy engagements across conditions

Condition	*M*	*SE*	LB	UB
Bubbles	0.28	0.01	0.26	0.30
Social	0.28	0.01	0.27	0.30
Non-social	0.29	0.01	0.28	0.30
Toy engagement	0.24	0.01	0.23	0.25

Table 4.17: Mean, standard deviation and N for average alpha power including all toy engagements across conditions in the frontal and posterior regions

Region	Condition	*M*	*SD*	*N*
Frontal	Bubbles	0.28	0.07	12
	Social	0.27	0.05	35
	Non-social	0.29	0.05	37
	Toy engagement	0.24	0.04	65
Posterior	Bubbles	0.28	0.06	12
	Social	0.29	0.07	35
	Non-social	0.29	0.06	37
	Toy engagement	0.24	0.04	65

Figure 4.6: Mean and confidence intervals for alpha power over different conditions when all toy engagements were included



4.3.2.2.2 Theta

An ANOVA of the linear regression found a significant main effect of condition on theta power (Table 4.18). Follow-up pairwise comparisons with Bonferonni correction found that theta power was lower during bubble blowing compared to the social interaction and toy engagement conditions (Table 4.19 and Table 4.20). Theta power during social interaction and toy engagement was also significantly higher than during the non-social condition (Table 4.19 and Table 4.20). Figure 4.7 displays these differences; Table 4.21 shows descriptive statistics for all conditions.

Table 4.18: **F**-value, p-value and partial eta squared effect size for ANOVAs on average theta power across regions and conditions when including all toy engagements (star indicates significance)

	F	*p*	η^2 (partial)
Region	0.03	.860	<0.01
Condition	8.37	< .0001*	0.08
Region x Condition	0.77	.512	0.01

Table 4.19: Estimate, t-value, p-value, and standard errors for pairwise comparisons average theta power (including all toy engagements) across conditions (star indicates significance)

		β	*t*	*p*	*SE*
Bubbles	Social	-0.05	-4.05	< .001*	0.01
Bubbles	Non-social	-0.02	-1.66	.347	0.01
Bubbles	Toy engagement	-0.04	-3.67	.002*	0.01
Social	Non-social	-0.03	-3.40	.004*	0.01
Social	Toy engagement	-0.01	-0.97	.769	0.01
Non-social	Toy engagement	0.02	2.92	.020*	0.01

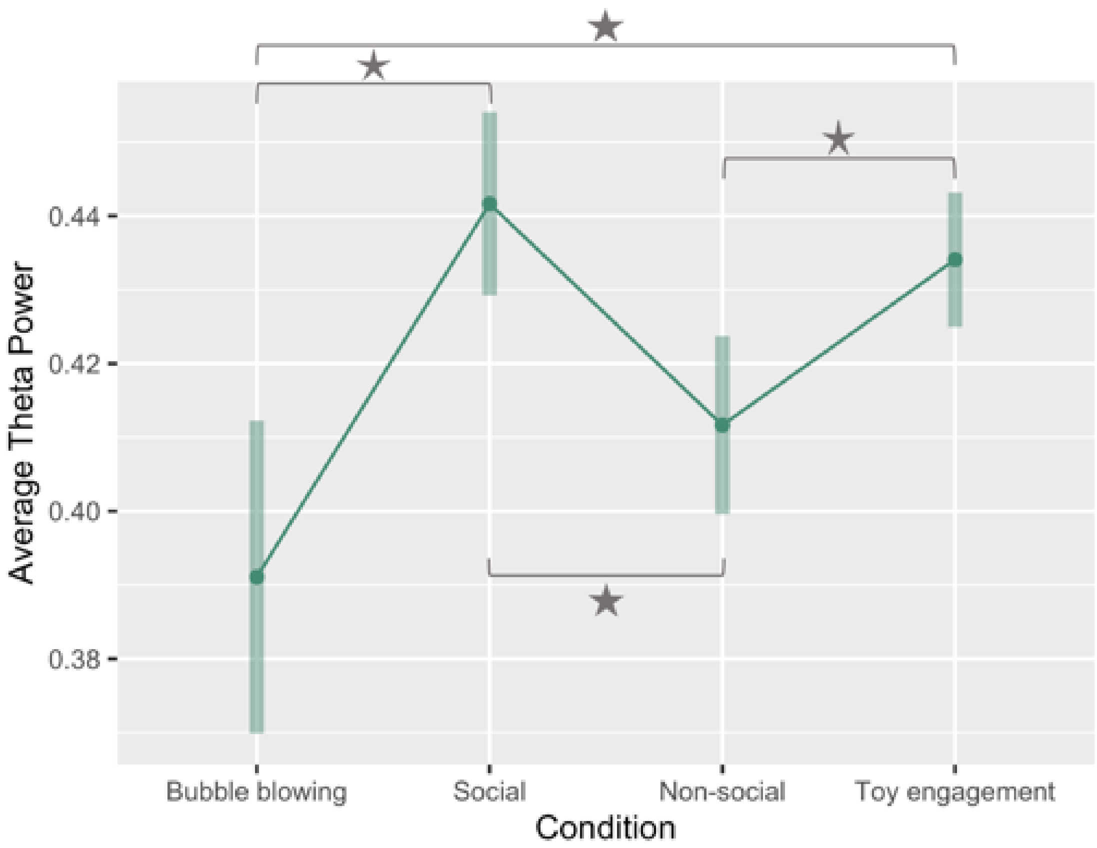
Table 4.20: Mean, standard error, and 95% confidence intervals for average theta power (including all toy engagements) across conditions

Condition	M	*SE*	LB	UB
Bubbles	0.39	0.01	0.37	0.41
Social	0.44	0.01	0.43	0.45
Non-social	0.41	0.01	0.40	0.42
Toy engagement	0.43	0.01	0.43	0.44

Table 4.21: Mean, standard deviation and N for average theta power (including all toy engagements) across conditions in the frontal and posterior regions

Region	Condition	*M*	*SD*	*N*
Frontal	Bubbles	0.39	0.05	12
	Social	0.44	0.06	35
	Non-social	0.41	0.05	37
	Toy engagement	0.44	0.05	65
Posterior	Bubbles	0.39	0.04	12
	Social	0.44	0.05	35
	Non-social	0.42	0.05	37
	Toy engagement	0.43	0.06	65

Figure 4.7: Mean and confidence intervals for theta power over conditions when all toy engagements were included



4.3.3 Change in EEG power during toy engagement

To assess whether there was any change in EEG power during engagement with toys, linear mixed models were fitted using EEG power as the dependent variable and segment number as a fixed effect. Participant was also included as a random effect allowing for both random slope and intercept, and engagement number was additionally included as a random effect allowing for random intercept to account for any potential differences across repetitions of toy engagement. Separate models were run for each of frontal and posterior alpha and theta power; only frontal alpha found a significant increase in power at the group level (4.3.3.1.1), but individual slopes and intercepts were extracted from each model for use in later analyses (4.3.4.3). Due to concerns about the assumption of normally distributed residuals, robust linear mixed models were conducted using *rlmer* in the robustlmm package in RStudio (Koller, 2016) for each model. The model was specified as below:

> *rlmer(Power ~ Segment number + (Segment number* | *Participant ID: Engagement number), data)*

Models were fitted using a 'nlminb' optimiser. A total of 1432 observations were included from 65 groups of participant and engagement number.

4.3.3.1 Alpha

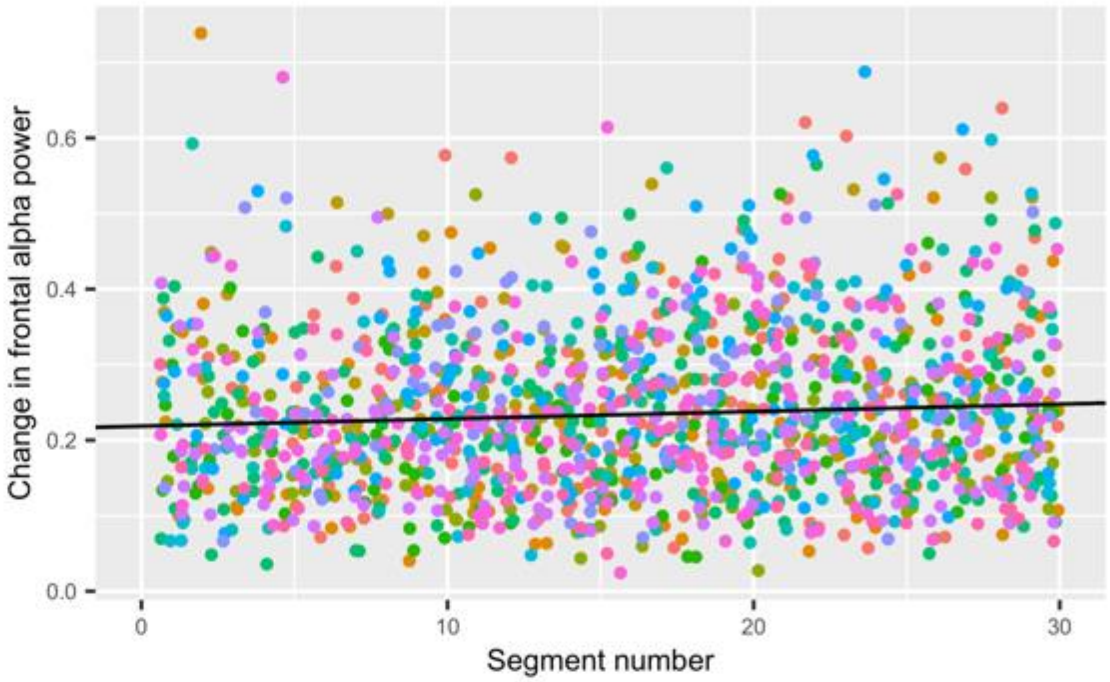
4.3.3.1.1 Frontal

The model fixed effect intercept was significant; β = 0.22, *SE* = 0.01, *t* = 30.88, $p < .001$, CIs[0.21, 0.32]; as was segment number; β = 0.001, *SE* < 0.001, *t* = 2.78, $p = .007$, CIs[<0.001, 0.002]; Figure 4.8. Random effects are shown in Table 4.22. The model fit was moderate; R^2(conditional) = 0.07, R^2(marginal) = 0.005.

Table 4.22: Variance and standard deviance for random effects for frontal alpha

	Variance	SD
ID	0.001	0.02
Segment number	<0.001	<0.001
Residual	0.011	0.11

Figure 4.8: Change in frontal alpha power over time during toy engagements



4.3.3.1.2 Posterior

The model fixed effect intercept was significant; β = 0.22, *SE* = 0.01, *t* = 26.75, *p* < .001, CIs [0.20, 0.24]; but segment number was not; β = 0.001, *SE* < 0.001, *t* = 1.93, *p* = .058, CIs [<-0.001, 0.002]. Random effects are shown in Table 4.23. The model fit was moderate; R^2(conditional) = 0.04, R^2(marginal) = 0.003.

Table 4.23: Variance and standard deviance for random effects for posterior alpha

	Variance	SD
ID	0.001	0.02
Segment number	<0.001	<0.001
Residual	0.016	0.13

4.3.3.2 Theta

4.3.3.2.1 Frontal

The model fixed effect intercept was significant; β = 0.43, *SE* = 0.01, *t* = 42.79, *p* < .001, CIs [0.41, 0.45]; but segment number was not; β = 0.001, *SE* = 0.001, *t* =

42.79, p = .137, CIs [<-0.001, 0.002]. Random effects are shown in Table 4.24. The model fit was moderate; R^2(conditional) = 0.10, R^2(marginal) = 0.002.

Table 4.24: Variance and standard deviance for random effects for frontal theta

	Variance	SD
ID: Engagement number	0.002	0.04
Segment number	<0.001	<0.001
Residual	0.020	0.14

4.3.3.2.2 Posterior

The model fixed effect intercept was significant; β = 0.42, *SE* = 0.01, t = 34.66, p < .001, CIs [0.39, 0.44]; but segment number was not; β = 0.001, *SE* = 0.001, t = 0.87, p = .391, CIs [<-0.001, 0.002]. Random effects are shown in Table 4.25. The model fit was moderate; R^2(conditional) = 0.07, R^2(marginal) = 0.001.

Table 4.25: Variance and standard deviance for random effects for posterior theta

	Variance	SD
ID	0.003	0.05
Segment number	<0.001	<0.001
Residual	0.027	0.17

4.3.4 Duration of toy engagement

Behavioural measures extracted from the free play session found that this was a feasible method of investigating exploration, with children providing a sufficient number and duration of toy engagements (4.3.4.1). Duration of toy engagements were further investigated in relation to neural measures collected during this period (4.3.4.2 and 4.3.4.3).

4.3.4.1 Free play session

Average length of free play session was around thirteen minutes. For thirteen of fourteen children, the free play session was at least ten minutes long with the session ending due to reaching the end of the research protocol rather than exceeding child tolerance. Most children engaged in play very quickly after entering the room, with ten of thirteen children engaging within fifteen seconds and only two taking longer than one minute. Of the 13 children for whom first engagement could be determined, 7 children engaged with puzzle 1 first. For five of these children the experiment was set up with the puzzle 1 at the size of the table furthest from where children entered the room, whilst this was reversed for the other two. Table 4.26 contains descriptive statistics for various free play measures.

Table 4.26: Descriptive statistics for measures relating to engagement during the free play session

	M	*SD*	LB	UB	*N*
Length of free play session	13m 8s	3m 10s	7m 44s	19m 32s	14
Time to first engagement (seconds)	20.0	25.9	3	90	13
Length of all engagements (seconds)	72.1	87.5	2	509	102
Length of first engagement (seconds)	41.5	42.9	2	158	13
Number of engagements per participant	7.6	2.7	4	15	14

4.3.4.2 Relation between average EEG power and length of engagement

To investigate whether length of toy engagement was related to average power, linear mixed models were performed. Power was included as the dependent variable with engagement length included as a fixed effect. As children provided data from multiple engagements, participant was included as a random effect with variable intercept. Engagement number (i.e whether it was the child’s first, second, third, etc. engagement) was additionally included as a random effect in a partially crossed design, in case there was a relation between this and average power. For some models, engagement number explained very little variance and reduced the model fit, therefore it was removed. Assumptions were checked, then models were run using the lmerTest package in RStudio (Kuznetsova et al., 2017) which applies Satterthwaite’s method to estimate degrees of freedom and generate p-values. The model specification was as follows:

lmer(power ~ engagement length + (1 | participant ID) + (1 | Engagement number), data)

Models included 59 observations from 13 participants and 10 different engagement numbers.

4.3.4.2.1 Alpha

4.3.4.2.1.1 Frontal

An ANOVA with Kenward-Roger's method of the linear mixed model of frontal alpha power found no significant effect of engagement length; $F(1,53.1) = 2.28$, $p = .137$ but the model fixed effect intercept was significant; $\beta = 0.23$, $SE = 0.01$, $t(31.29) = 22.21$, $p < .0001$; CIs [0.22, 0.25]. The random effect of participant explained around 28% of the variation in theta power that was not explained by fixed effects; engagement number explained approximately 5%; Table 4.27. The model fit was moderate; AIC = -191.67, BIC = -90.53, R^2(conditional) = 0.36, R^2(marginal) = 0.04, ICC = 0.33, RMSE = 0.03.

Table 4.27: Random effects from mixed model of relation between average frontal alpha and engagement length

	Variance	SD	LB	UB
ID	<0.001	0.02	0.01	0.03
Engagement number	<0.001	0.01	0.00	0.02
Residual	0.001	0.03	0.03	0.04

4.3.4.2.1.2 Posterior

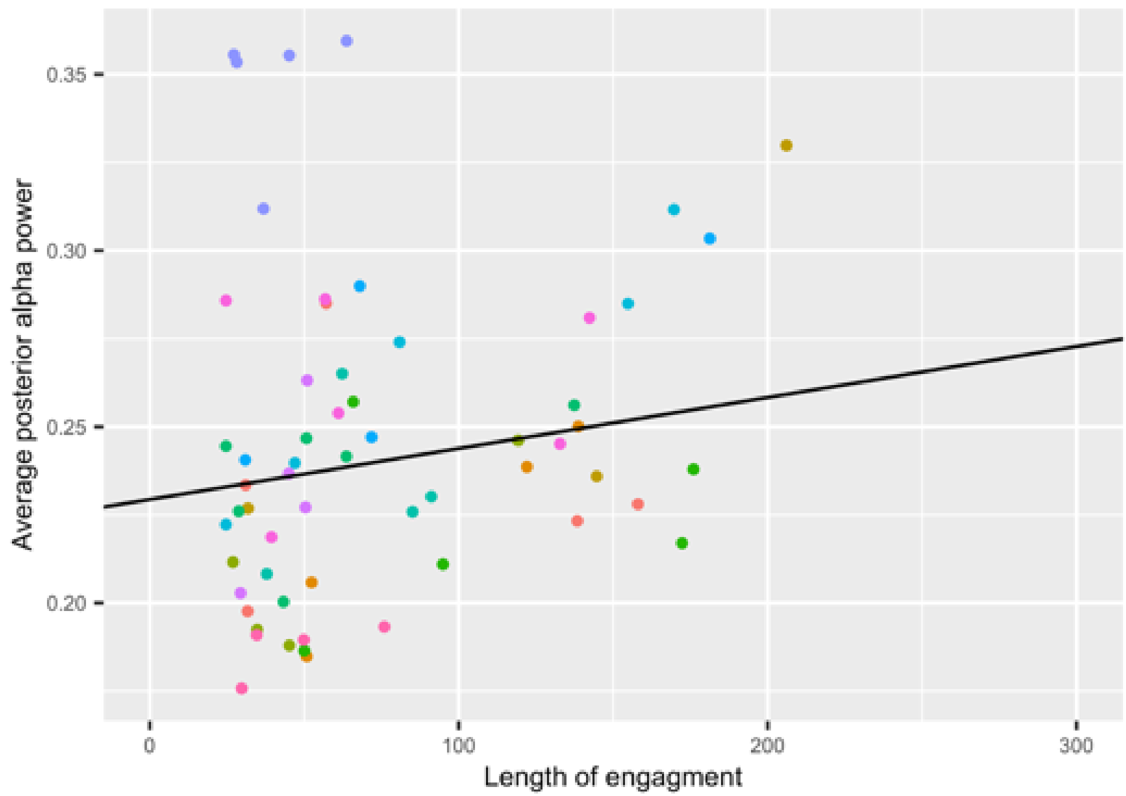
An ANOVA with Kenward-Roger's method of the linear mixed model of occipital theta power found a significant effect of engagement length; $F(1,47.43) = 15.51$, $p = .0003$; with estimates indicating this was in a positive direction, with longer engagement lengths associated with greater power; $\beta < 0.0003$, $SE < .001$, $t(47.6) = 4.10$, $p = .0002$, CIs[0.0001,0.0004]; Figure 4.9. The intercept was also significant; $\beta = 0.22$, $SE = 0.01$, $t(17.70) = 28.25$, $p < .0001$; CIs [0.21, 0.25]. The random effect of participant explained around 71% of the variation in theta power that was not explained by fixed effects; engagement number explained approximately 1%; Table

4.28. The model fit was moderate; AIC = -205.49, BIC = -195.10, R^2(conditional) = 0.75, R^2(marginal) = 0.10, ICC = 0.72, RMSE = 0.02.

Table 4.28: Random effects from mixed model of relation between average posterior alpha and engagement length

	Variance	SD	LB	UB
ID	0.002	0.04	0.03	0.06
Engagement number	<0.001	0.004	0.00	0.01
Residual	0.001	0.02	0.02	0.03

Figure 4.9: Relation between length of engagement and average posterior alpha power (different participants in different colours)



4.3.4.2.2 Theta

4.3.4.2.2.1 Frontal

A linear mixed model was fitted as outlined using average frontal theta power as the dependent variable. Assumption checking revealed concerns about the normality of residuals for this model via visual inspection, but Shapiro Wilk normality test indicated normality could be assumed (W = 0.97369, p = 0.2292), therefore this analysis was used. When included in the model, engagement number explained close to zero variance and reduced the model fit, therefore it was excluded from the model. The model fixed effect intercept was significant; β < 0.44, *SE* < .002, *t*(18.20) = 29.66, $p < .0001$, CIs [0.42, 0.49], but an ANOVA of the linear mixed model using Kenward-Roger's method found no evidence of a significant effect of engagement length; $F(1,47.5) = 0.07$, $p = .790$. The random effect of participant explained around 70% of the variation in theta power that was not explained by fixed effects; Table 4.29. The model fit was moderate; AIC = -185.42, BIC = -177.11, R^2(conditional) = 0.70, R^2(marginal) = 0.001, ICC = 0.70, RMSE = 0.03.

Table 4.29: Random effects from mixed model of relation between average frontal theta and engagement length

	Variance	SD	LB	UB
ID	0.002	0.04	0.03	0.06
Residual	0.001	0.03	0.03	0.03

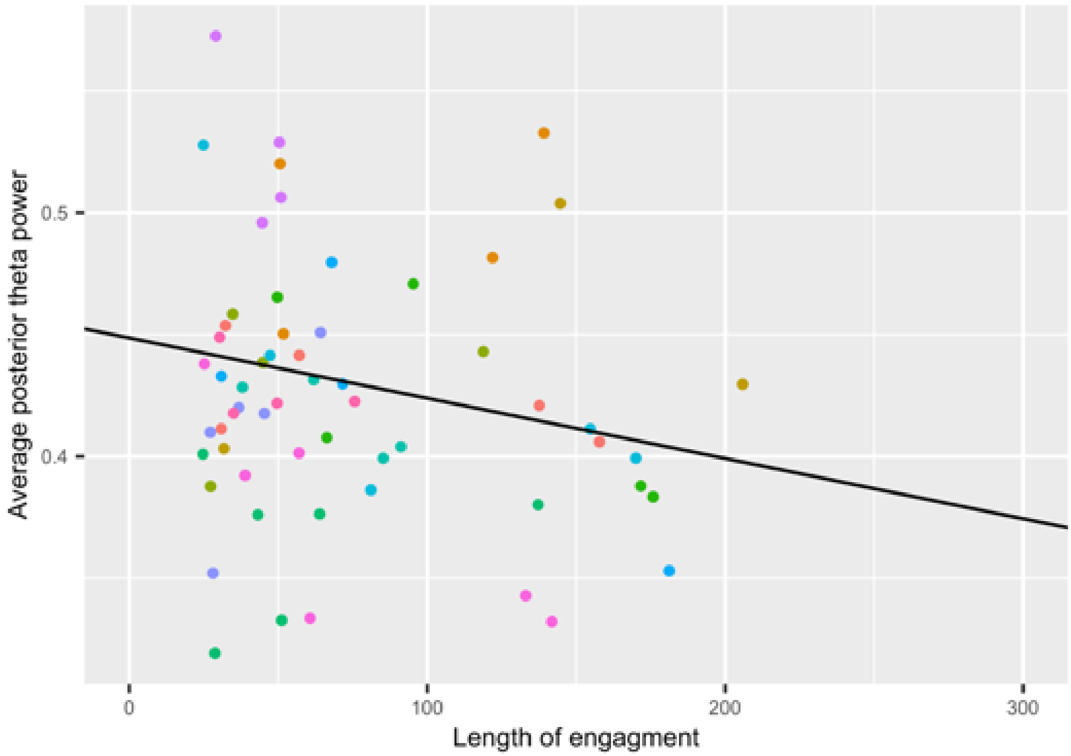
4.3.4.2.2.2 Posterior

An ANOVA with Kenward-Roger's method of the linear mixed model of occipital theta power found a significant effect of engagement length; $F(1,49.3) = 5.29$, $p = .026$; with estimates indicating this was in a negative direction, with shorter engagement lengths associated with greater power; β < -.0002, *SE* < .001, $t(49.4) = -2.32$, $p = .024$, CIs[-.0005,-.0001]; Figure 4.10. The intercept was also significant; β = 0.45, *SE* = 0.02, $t(22.82) = 30.71$, $p < .0001$; CIs [0.43, 0.50]. The random effect of participant explained around 55% of the variation in theta power that was not explained by fixed effects; engagement number was removed from this model as it explained approximately 0 variance and reduced model fit; Table 4.30. The model fit was moderate; AIC = -170.22, BIC = -161.91, R^2(conditional) = 0.58, R^2(marginal) = 0.05, ICC = 0.55, RMSE = 0.03.

Table 4.30: Random effects from mixed model of relation between average posterior theta and engagement length

	Variance	SD	LB	UB
ID	0.002	0.04	0.02	0.05
Residual	0.001	0.04	0.03	0.04

Figure 4.10: Relation between length of engagement and average posterior theta power (different participants in different colours)



4.3.4.3 Relation between EEG power change and length of engagement

EEG power change measured the association between power and time over the course of each bout of toy engagement; length of engagement was the time (in secs) that same toy engagement lasted (see 4.3.3 for details). To investigate whether length of toy engagement was related to change in power, linear mixed models were performed. Power change extracted from linear mixed models (4.3.3) was included as the dependent variable with engagement length included as a fixed effect. Power change coefficients were first multiplied by 100 before use in these models to avoid scaling issues in the current analyses. Models were otherwise the same as in section 4.3.4.2, and again included 58 observations from 10 engagement numbers and 13 participants.

4.3.4.3.1 *Alpha*

4.3.4.3.1.1 *Frontal*

An ANOVA of the linear mixed model of occipital theta power found a non-significant effect of engagement length; $F(1, 52.31) = 0.74$, $p = .394$. The intercept was not found to be significant; β = 0.004, $SE = 0.01$, $t(29.95) = -0.78$, $p = .444$; CIs [-0.01, 0.01]. The random effect of participant explained very little variation in alpha power that was not explained by fixed effects; Table 4.31. The model fit was moderate; AIC = -264.47, BIC = -254.17, R^2(conditional) = 0.35, R^2(marginal) = 0.01, ICC = 0.34, RMSE = 0.01.

Table 4.31: Random effects from mixed model of relation between frontal alpha change and engagement length

	Variance	SD	LB	UB
ID	<0.001	0.01	0.01	0.02
Engagement number	<0.001	0.01	0.00	0.01
Residual	<0.001	0.02	0.01	0.02

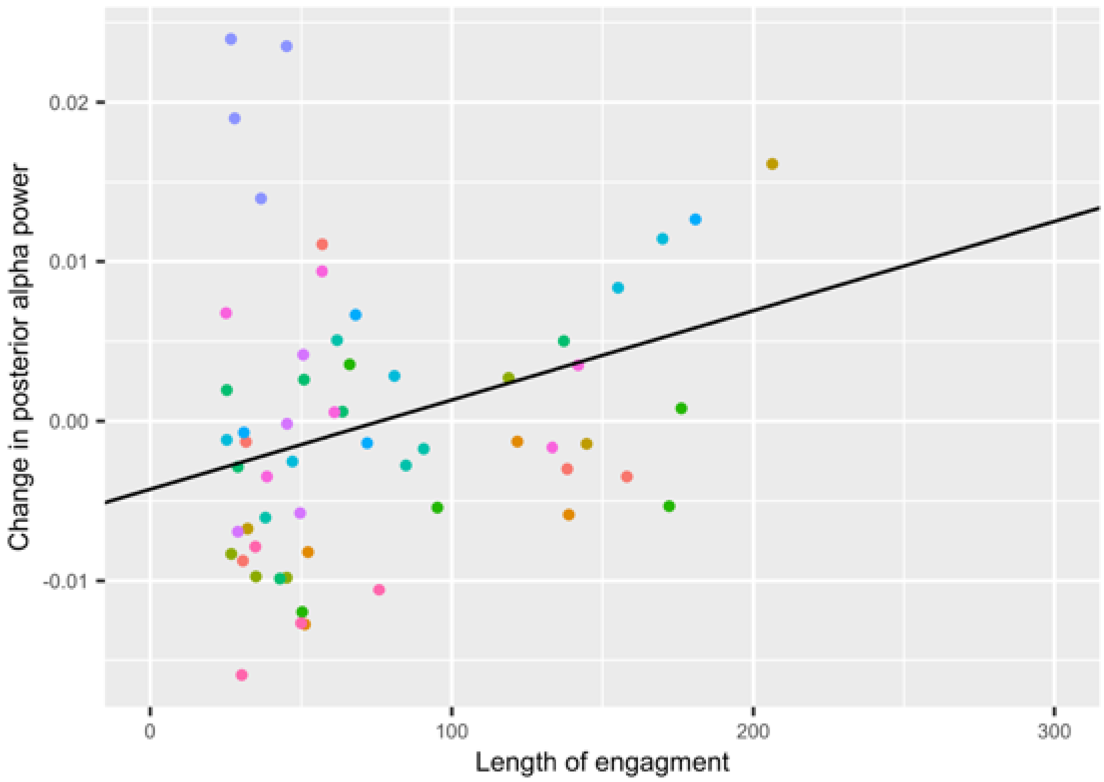
4.3.4.3.1.2 *Posterior*

An ANOVA with Kenward-Roger's method of the linear mixed model of occipital theta power found a significant effect of engagement length; $F(1,46.57) = 12.63$, $p < .001$, with the estimate indicating this was small and positive; β < 0.001, $SE < 0.001$, $t(46.57) = 3.55$, $p < .001$; CIs<0.001, <0.001]; Figure 4.11. The intercept was not significant; $\beta = -0.004$, $SE = 0.002$, $t(18.64) = -1.71$, $p = .104$; CIs [-0.007, 0.001]. The model fit was moderate; AIC = -374.99, BIC = -364.68, R^2(conditional) = 0.70, R^2(marginal) = 0.09, ICC = 0.67, RMSE = 0.01. Random effects of ID and engagement number explained very little variance; Table 4.32.

Table 4.32: Random effects from mixed model of relation between posterior alpha change and engagement length

	Variance	SD	LB	UB
ID	<0.001	0.01	0.01	0.01
Engagement number	<0.001	<0.01	0.00	0.002
Residual	<0.001	0.01	0.004	0.01

Figure 4.11: Relation between length of engagement and change in posterior alpha power (different participants in different colours)



4.3.4.3.2 Theta

4.3.4.3.2.1 Frontal

An ANOVA of the linear mixed model of frontal theta power found a non-significant effect of engagement length; $F(1, 46.04) = 0.01$, $p = .918$. The intercept was not found to be significant; $\beta < 0.001$, $SE < 0.01$, $t(16.88) = 0.18$, $p = .862$; CIs [-0.01, 0.02]. The random effect of participant explained very little variation in alpha power that was not explained by fixed effects; Table 4.33. The model fit was moderate; AIC = -272.06, BIC = -261.76, R^2(marginal) < 0.001, RMSE = 0.01.

Table 4.33: Random effects from mixed model of relation between frontal theta change and engagement length

	Variance	SD	LB	UB
ID	0.001	0.02	0.01	0.03
Engagement number	<0.001	<0.001	0.00	0.002
Residual	<0.001	0.01	0.01	0.02

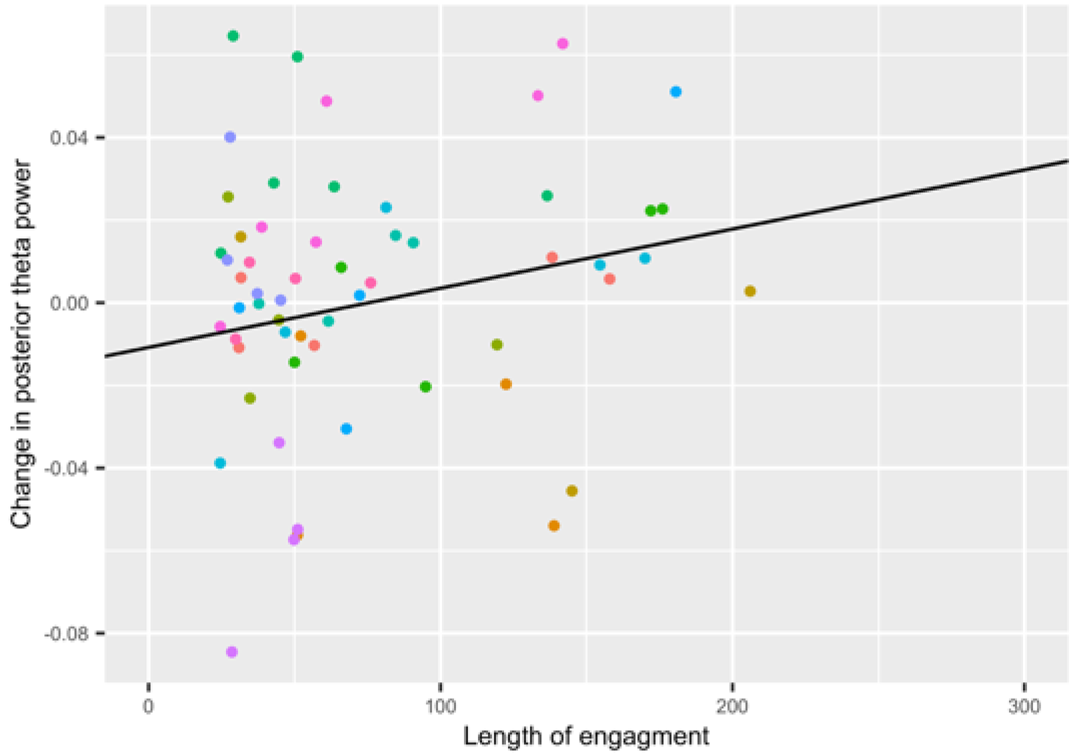
4.3.4.3.2.2 Posterior

An ANOVA of the linear mixed model of occipital theta power found a significant effect of engagement length; $F(1, 48.54) = 5.54$, $p = .023$, with the estimate indicating this was small and in a positive direction; $\beta < 0.001$, $SE < 0.01$, $t(48.54) = 2.35$, $p = .0.23$; CIs[<0.001, 0.002]; Figure 4.12. The intercept was not significant; $\beta = 0.011$, $SE = 0.01$, $t(22.70) = -1.30$, $p = .208$; CIs [-0.018, 0.005]. The model fit was moderate; AIC = -227.46, BIC = -217.16, R^2(marginal) = 0.11, RMSE = 0.02. Random effects of ID and engagement number explained very little variance; Table 4.34.

Table 4.34: Random effects from mixed model of relation between posterior theta change and engagement length

	Variance	SD	LB	UB
ID	0.001	0.02	0.02	0.03
Engagement number	<0.001	<0.001	0.00	0.002
Residual	<0.001	0.02	.002	0.02

Figure 4.12: Relation between length of engagement and change in posterior theta power (different participants in different colours)



4.3.4.4 Exploratory behaviours in relation to anxiety and environmental measures

To investigate whether there were any associations between anxiety and exploration style, and environment and exploration style, correlations were performed. Mean engagement length was calculated for each participant and was used as a measure of exploratory style. Table 4.35 shows descriptive statistics for these variables from the GAD-7, BIS, neighbourhood scale and home-CHAOS questionnaire, as well as mean engagement length per participant.

Table 4.35: Descriptive statistics for anxiety, environmental and exploration measures

	Recruited sample				
	M	*SD*	Min.	Max.	*N*
Parent GAD score	4.7	4.3	0	11	16
Child BIS score	16.4	3.3	12	24	16
Neighbourhood score	41.1	21.4	12	77	16
CHAOS score	27.2	7.1	20	41	16
Mean engagement length	1min 23secs	43.2secs	26.5secs	178.8secs	13

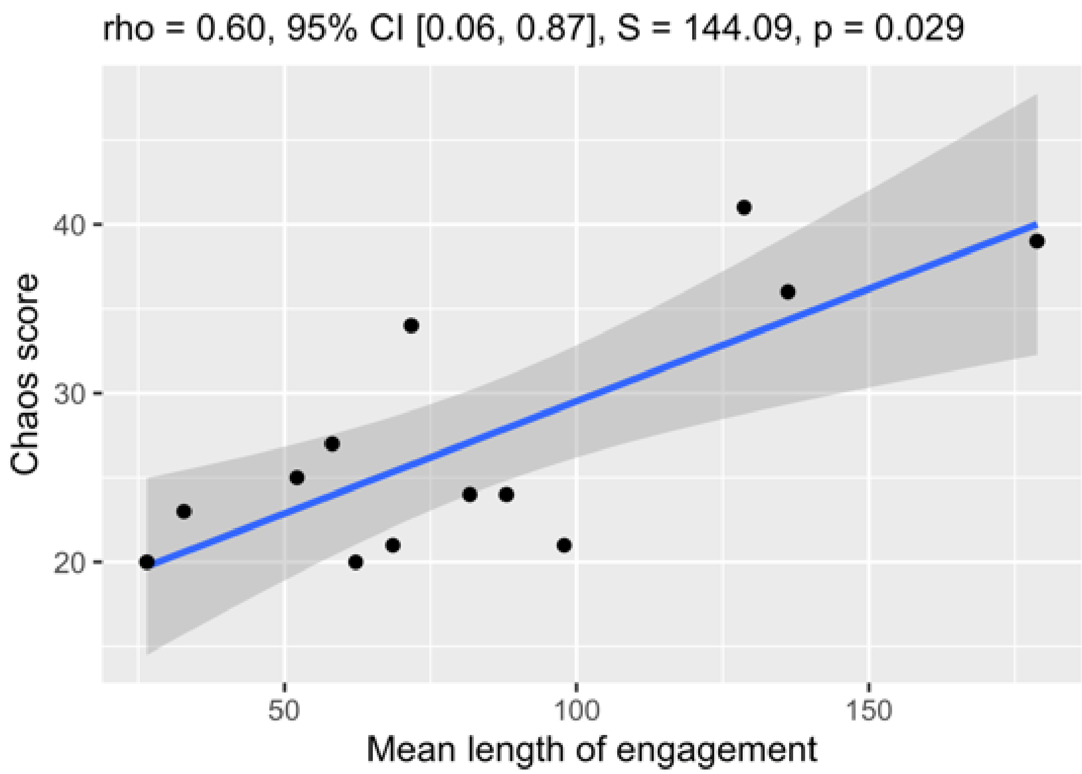
4.3.4.4.1 Correlations with anxiety measures

Parent GAD score was not normally distributed (W = 0.86, p = .04), therefore Spearman's rank was used for correlations involving GAD score, whilst Pearson's correlations were used for BIS score. The association between GAD and average length of engagement per participant was not found to be significant; $r(11)$ = 0.45, p = .123; the association with BIS score and average length of engagement was also not significant; $r(11)$ = -0.13, p = .669.

4.3.4.4.2 Correlations with environmental measures

CHAOS score was not normally distributed (W = 0.84, p = .02), therefore Spearman's rank was used for correlations involving CHAOS score, whilst Pearson's correlations were used for Neighbourhood score. The association between CHAOS score and average length of engagement per participant was positive and significant at the $p<.05$ level but not after Bonferonni corrections for multiple comparisons were applied; $r(11)$ = 0.60, p = .029; Figure 4.13; the association with BIS score and average length of engagement was not significant; $r(11)$ = 0.70, p = .250.

Figure 4.13: Relation between mean length of engagement and chaos score



4.4 DISCUSSION

The current chapter assessed the feasibility of a naturalistic design for collecting electroencephalography data from 2- and 3-year-olds and considered differences in EEG relating to conditions engaging different cognitive processes.

4.4.1 Feasibility of less-controlled design

Feasibility checks found that a good amount of valid data was collected in all EEG conditions and parents rated the study highly, indicating the suitability of the design for this purpose. Slightly fewer segments collected during third repetitions of social and non-social interactions were usable than during first and second trials. This may reflect a level of disengagement from the condition likely due to boredom. Whilst enough usable data was collected for the current analyses, it might be useful to adapt the protocol to prevent this reduction in usable segments such that differences across repetitions can be validly investigated. The smaller number of usable segments reported for free play engagement is due to a shorter time period for this condition (30 versus 60 seconds) and is analogous to other conditions when multiplied to a similar length. These results suggest that EEG data can be collected during this free play design that is of similar quality and quantity to that which is collected during traditional designs in which children are seated. This could be of particular use with toddlers who have limited attention abilities and may prefer to physically explore their environment, whilst additionally improving ecological validity of measures. Such designs are facilitated by the development of portable EEG systems and the current findings indicate the suitability of the Enobio EEG system (NE, Neuroelectrics, Barcelona, Spain) for this purpose.

In addition to feasibility analyses of EEG data, questionnaire data collected from parents indicated high acceptability of the current study design. On a range of questions relating to practical and theoretical elements of the study, no parents responded with less than four out of five on a scale of not satisfied (1) to very satisfied (5). Importantly, this included questions about the assessment of their child and the use of this in assessing neurodevelopment from a parent's point of view. Such high satisfaction indicated that parents were comfortable with the study protocol and considered it suitable for their toddler-aged children, further supporting high feasibility of this less-controlled design for neurocognitive research with toddlers.

4.4.2 Exploration and brain activity

Conditional analyses revealed different patterns of theta and alpha activity across conditions. As predicted and similar to other work (Orekhova et al., 2006), average theta power was higher in the social and toy engagement conditions compared to both non-social and bubble blowing conditions. This makes sense when considered in light of work which implicates theta in active learning and may contribute to a wider role for theta in establishing optimal opportunities for learning (Begus & Bonawitz, 2020). Similar patterns were found involving all engagements and initial engagements only, though significance was not found for the latter, likely due to lack of statistical power caused by small sample size. In addition to condition comparisons, it was further predicted that there may be a positive association between theta power and depth of exploration (i.e. length of a toy engagement in seconds; see 4.2.4.2). This prediction was based on suggestions that theta activity might predict how long children explore for (Begus & Bonawitz, 2020) and findings that showed theta power preceding visual fixations predicted duration of subsequent fixation in 12-month-olds (Wass et al., 2018). Analyses found no significant association between average frontal theta during a toy engagement and depth (length) of that exploration bout, whilst higher average posterior theta during a bout was found to be related to shorter exploration bouts. These findings may seem puzzling given that higher theta was found during exploration at the condition level, however they suggest that, whilst exploration may involve greater theta power than other activities, this does not translate to depth of exploration at the individual level. Such findings highlight the difference between analyses which average over a condition versus those which consider power at the moment-to-moment level and emphasise the importance of investigating both. They further support a growing body of evidence which indicates that variability in power over a short period (i.e. at the moment-to-moment level) might be uniquely informative about neural processing (Garrett et al., 2013) and emphasise the need for future work to further explore these measures.

Intriguingly, different relations with depth of exploration were found for average theta and theta change measures. Whilst a negative relation was found between depth of exploration and average theta power, a positive relation was found between exploration and theta change in the posterior region. This suggests that longer engagements were associated with lower average theta but a greater degree of change (either increase or decrease) and suggests consideration of both measures may provide useful information about underlying processing. Further, firm predictions were not made about regional differences and condition comparisons did not show significant differences in power between frontal and posterior regions, but the fact that these findings are specific to the posterior region may provide insight into how activity is distributed across the brain. In fact, this finding appears to contradict work which found predominantly frontal theta increases during exploratory behaviours in infant and preschoolers, but may be due to differences in experimental design, measures used or age of participants. Future work may benefit from exploring the topography of theta and alpha power during exploratory behaviours in a longitudinal design, such that changes in this can be better understood.

Analogous analyses involving alpha power also revealed positive associations with depth of exploration which were apparent only in posterior regions, with the same pattern found for average alpha and alpha change measures. This conflicts with the hypotheses, which predicted a negative association between alpha power and exploration and is seemingly in contrast to the finding that average alpha power was significantly lower during toy engagement compared to all other conditions, with no significant regional differences found. The latter finding was in line with predictions in the current chapter which were largely based upon work indicating that higher alpha power occurs in relation to instances of focussed attention (Orekhova et al., 2001). This has been interpreted as indicating that alpha may be important for inhibition of attention to other stimuli (Klimesch et al., 2007), whilst distributed attention has been outlined as crucial for enhanced exploration (Blanco & Sloutsky, 2020). In contrast to exploratory behaviours engaging a broader attention strategy, bubble blowing, social and non-social conditions likely involved more focussed attention, hence may require greater inhibition of attention to wider visual fields and involve higher alpha activity. The current findings of lower alpha power during the free play condition which involves more widely distributed attention might therefore be taken as further support for a role of alpha in inhibitory control, though further research is required to determine if this is the case, particularly given the positive association between alpha power and exploration length as discussed above.

It is possible that a positive relation between posterior alpha power and exploration might be explained by the choice of measure for exploration, which considered the length of individual bouts of toy engagement. Though the free-play exploration session might generally engage distributed attention, focussed attention may be required during specific bouts of toy engagement. That higher alpha was found during longer bouts of engagement may be reflective of a change from distributed to focussed attention which occurs to a greater degree when children are focussed on an activity for a longer period. Whilst the current study might thus suggest that there was an aspect of the exploration condition that induced lower alpha than other conditions, that a positive association between alpha and exploration was found suggests this may not be driven by lower alpha during toy engagements. Future research which uses other measures of exploration may add to understanding about the role that alpha plays. Nonetheless, the current findings do provide empirical evidence of lower alpha measured during a free-play session compared to during bubble blowing, a social and non-social condition in toddlers, whilst the opposing findings demonstrate the use of different approaches (i.e. condition differences and relations between measures) in investigating neural functioning.

4.4.3 Exploration and early experiences

Based upon findings which indicated that experience of adversity and higher anxiety were related to reduced exploration in individuals who had previously experienced institutionalisation (Humphreys et al., 2015), the current study sought to investigate this relation in a sample who had experienced less extreme adversity. Interestingly, these analyses indicated that children from more chaotic homes (as indexed by higher scores on the home CHAOS questionnaire) tended to have longer bouts of toy exploration, though this finding should be taken with caution given that significance was not reached after correction for multiple comparisons were applied. If we consider that a higher CHAOS score indicates greater confusion and disorganisation of a home (which may be characteristics linked to higher adversity), it may be counterintuitive to expect a negative relation, given that greater exploration may be advantageous in more unpredictable contexts (i.e. where there is greater confusion). Specifically, chaotic, and unpredictable environments may provide fewer opportunities for engagement, meaning children may learn to actively seek these out. In addition, if greater exploration is considered to be related to more distributed attention, this positive relation between experience of a disorganised environment and exploration might also be linked to a broader field of attention differences in children from different backgrounds, as discussed in chapter 2 in this thesis. That significance was not reached in the current study could potentially be explained by a lack of power due to limited sample size (n = 13), meaning that it is difficult to draw conclusions at this stage, though it is still possible that a different pattern may emerge with a larger sample.

4.4.4 Limitations and future direction

An obvious limitation of the current work is the limited sample size. Where possible, analyses involved multiple trials per participant and utilised models which maximised data inclusion (such as linear models), however correlations between depth of exploration and environmental and anxiety measures may be particularly underpowered. Findings should thus be interpreted with caution and future work should focus on repeating analyses with a larger sample size. Nonetheless, the main aim of the current chapter was to assess the feasibility of a less-controlled design for collecting neurocognitive data with toddlers and findings can be interpreted for this purpose, and this work considered a blueprint for future research.

When considering how feasible the current study design is, it is of note that any system which involves wearables will likely result in some level of refusal from young children; indeed, in the current study two children refused to wear the EEG cap. One further child completed the bubble blowing, social and non-social interaction conditions but did not engage with activities during free play; whilst this did not appear to be due to shyness, it may have been due to child temperament, cognitive ability, interest in the activities or something about the experiment setting. It is hoped that conducting research in other settings with which children are more comfortable may reduce the impact of such factors on how children perform, however individual differences will always remain between how children engage with a study. Future work may consider altering the included table-top activities to include objects that engage a wider range of children and adjusting protocols for researcher/ parent engagement during the free play session to support all children's engagement whilst retaining a level of experimental control between participants.

It may also be beneficial to consider other measures of exploratory behaviour in future work of this kind. The current study used time of engagement as the main measure of exploration, which was based upon the depth of exploration measure used by van Liempd et al. (2018) but calculation differed slightly. Whilst this has revealed different associations with theta and alpha power, other measures of exploration (i.e. grouping by exploratory style or use of other behavioural coding schemas) may enable this relation to be explored more deeply. While such analyses require greater sample sizes than provided here, the current study design does allow for other measures to be used.

An additional consideration for less-controlled designs such as this, is the potential for motion to have an impact on the observed results. During bubble blowing, social and non-social interactions children remained seated, whereas toy exploration data was gathered during free play, when children were free to move around a room. It is likely, therefore, that children moved more during the exploration compared to other conditions. Given that greater motion causes higher artifact contamination (Gorjan et al., 2022), it is possible that the current findings of differences in power during toy exploration were confounded by more motion in that condition. Of note, the same procedures for data processing were applied across conditions and all artifact identification was done as one, therefore it is likely that movement artifacts were equivalently removed from all EEG data. Nonetheless, future work that explores and controls for degree of movement in its analyses would ensure that the observed condition differences could not be explained by motion differences alone. In addition, the careful choice and application of automatic methods for artifact detection and removal might help to minimise any potential impact of movement on EEG findings (Gorjan et al., 2022).

4.4.5 Conclusion

The current chapter focusses on a naturalistic study design to collect EEG data from a sample of 2- and 3-year-olds. This design was motivated by a need to adapt methods to improve neurocognitive data collection from toddlers and expand research into less-controlled settings (Bhavnani et al., 2021) such as in communities or homes. Toddlerhood constitutes a period of significant development which signals it as an important age to study whilst adding additional challenges for researchers in handling children's motor and attention abilities. The current study found that a free play session in which children could play with table-top activities facilitated collection of good quality and quantity of EEG data that was equivalent to more traditional conditions in which children were seated and watched a set stimulus. Condition comparisons found expected patterns of theta and alpha power, further supporting a role for theta in active learning (Begus & Bonawitz, 2020) and alpha in inhibitory control (Klimesch et al., 2007; Orekhova et al., 2001). Relations were also found between depth of exploration and each of theta and alpha in posterior regions, though further investigation is needed to fully understand these findings. A relation was also found between experience of chaotic environments and depth of exploration, which may be considered as a useful adaptation based on context and could be explored in relation to hypotheses about experience-related attentional differences. In addition to finding high feasibility for this less-controlled experimental design using a portable EEG system with 2- and 3-year-olds, this chapter contributes empirical findings about theta and alpha activity in toddlers.

5 APP-BASED TOOL FOR REMOTE DATA COLLECTION

Abstract

A considerable body of evidence has indicated a relation between children's early experiences and cognitive development, however much developmental cognitive neuroscience research is severely lacking in representative participant samples. This means findings may be missing details about the full nature of experience-development associations and are limited in their generalisability. Reasons for a lack of diversity in this research are likely multifaceted and complex, however difficulties associated with attending research settings and attitudes or perceptions about research organisations may play a role. The current chapter sought to investigate methods for increasing the diversity of participants in developmental research, through the development of a scalable app-based measure of early development. Such a tool may be used by researchers to remotely collect developmental data about young children, thereby reducing some burdens traditionally associated with participation in developmental cognitive neuroscience (DCN) research, and could ultimately be utilised to better understand the relation between early experiences and cognitive development. Focus groups and questionnaire data were used to identify factors which parents consider important for research utilising an app-based tool, which could have influential implications for future development of this research. A current app-based tool revealed strong relations between app measures, age, and other cognitive measures, supporting the validity of this tool for cognitive data collection. Finally, a data-driven approach found two clusters among numerous SES variables which mapped to a low and high SES group as is typically used in research, though analyses did not reveal any relations between SES grouping and cognitive ability. These findings are together informative about methods for improving the diversity of representation in future developmental cognitive research and provide an improved understanding of how an app-based tool might be used to improve diversity in DCN research. This could have important implications for future research investigating the relation between children's early experiences and cognitive development.

Situate in thesis

Whilst the previous chapter was focussed on the development of a neurocognitive paradigm which could be used across settings, the current chapter focussed on the development of a scalable digital tool which can be used for completely remote data collection. A digital tool that can be used remotely may reduce some practical barriers associated with taking part in studies at a research or other settings and might facilitate two-way information sharing which could, over time, help to create an equal partnership between researchers and participants. This chapter used focus groups to engage parents and to gain understanding about their views of research, child development, and an app-based digital tool. It also assessed the validity of an existing app for collecting cognitive data from infants. Given the potential for this app-based tool to help increase diversity of participation in DCN research, and the first step towards engaging communities in research afforded by focus groups, this chapter fits with the aims of the overall thesis by developing a tool and expertise for increased diversity and representation of participant samples in developmental cognitive neuroscience (DCN) research.

5.1 INTRODUCTION

Developmental cognitive neuroscience research has made a wealth of findings about children's early development, though this information is not always effectively disseminated to people who interact daily with young children, namely parents and childcare workers. One key finding is the impact that early experiences can have on an individual's development and success throughout life (Bradley & Corwyn, 2002; Ursache & Noble, 2016). Remarkably, these findings emerge despite a lack of diversity of experiences of families typically included in developmental cognitive neuroscience research. It may be considered that both these issues are flipsides of the same coin; that is, they both relate to challenges associated with engaging a broader audience in research. One way to overcome these difficulties may be to develop a remote, scalable measure of early development which suits both researcher and parent needs.

Though there is an increasing interest in scalable methods, a tradition of studies involving small sample sizes collected in highly controlled laboratory or healthcare settings means most developmental knowledge is still based upon such work. Such studies are particularly limiting as they typically include biased representation, with children of high socioeconomic status (SES) families typically overrepresented compared to low SES families (see Green et al., 2022). Not only does this limit the generalisability of findings from these studies, but some authors have suggested it also seriously limits external validity (Mulder et al., 2014). Whilst many psychology researchers have recognised the need to improve diversity of representation in studies, the practicalities of this are harder.

5.1.1 Factors affecting engagement in research

Factors influencing whether individuals engage in research are likely multifaceted and may not always be recognised by the individual themselves, whilst researchers' own position and perspectives might influence their role in recruiting participants. Some barriers such as potential difficulties with travel requirements, childcare arrangements, monetary costs, or language differences may be more tangible and explicit, but are not necessarily more influential, than others. For example, individual differences in attitudes, perceptions, knowledge, and beliefs may have a significant influence on an individual's engagement in research (Garcini et al., 2022). Different individuals may have varying levels of trust in academic and research institutions, which may have been perpetuated by historical interactions with researchers or organisations. For instance, people from communities which have been historically marginalised or under-represented may have particularly high levels of mistrust in organisations, which could influence their likelihood of engaging in research. Mistrust or apprehension to take part in research may also relate to concerns about misuse of personal data (Garcini et al., 2022). Other individuals may lack knowledge or have misconceptions about what research is and why it is important, which may cause them not to engage. Furthermore, there may be existing processes or systems which prevent or exclude certain individuals from partaking in research as well as systematic biases in the way research is approached, which may form a significant hurdle for some individuals. As outlined by Garcini et al. (2022), barriers at systematic and structural levels could refer to "attitudes, rules, regulations, policies and structures within research institutions and healthcare systems that may assist or hinder research participation or research of certain topics" (p7). Thus, despite this not being an exhaustive list, it is already clear that there is a large variety of factors which might influence an individuals' likelihood to participate in a research study and which developmental cognitive research must attempt to overcome.

5.1.2 Methods for making research more accessible

Some difficulties may be improved by practical changes which are relatively easy for researchers to implement. For example, covering travel costs upfront, reimbursing families for their time, and offering childcare options, could significantly reduce burdens for some families. Other organisational elements such as having flexible schedules which fit around families' commitments, providing easy-to-understand information and documents, and tailoring resources to participants (i.e. by translating to different languages, providing visual and written forms) might also improve the accessibility of research to a wider range of individuals (Garcini et al., 2022). Taking time to explain the whole research process to participants early in the study and ensuring researchers are receptive and responsive to answering questions, as well as actively support arrangements, could also help to build trust between research institutes and families, and ensure participants have a positive experience. Such strategies require some extra work by researchers but could be relatively easily achieved and may have a substantial impact in helping more families take part in research.

Other methods to increase diversity in developmental research might require much more considerable changes in approach. A shift from laboratory-based studies to research which is moveable and can be carried out in communities may be one such method. This approach could reduce practical barriers such as those associated with travel and childcare, as well as potentially increasing trust, by empowering communities to play a larger role in the organisation of research and enabling an element of familiarity with the research setting. Community-based research has been facilitated by technological advancements such as the development of portable neuroimaging systems and developments with smart phones and mobile apps. The latter additionally provides options for completely remote research to be conducted, without a need for participants to physically visit any setting. Though this method likely would not reduce all barriers (families may still have concerns about data management, etc.), it could provide a method which is accessible for a much broader range of families. Indeed, a majority of the World's adult population own a mobile device (Phillips et al., 2024), meaning app-based technologies hold potential to reduce some socioeconomic disparities associated with access to information around early child development (Crouse et al., 2023). In addition to broadening participation, app-based methods are cheap and easy to scale, meaning they could additionally improve the size of studies and power of research findings.

As well as a shift from laboratory-based research, involving community members in research could also increase trust between participants and researchers, and help researchers make studies more accessible for particular communities. Community engagement - meaningfully involving members of a community in the research process (Han et al., 2021) - may help researchers to understand factors that are impacting that community specifically, which they may not otherwise be aware of. Participants from different backgrounds may have differing preferences in cultural values, communication and interpersonal styles which crucially may be different to researchers' own and which could have an important impact on an individual's experience of research (Garcini et al., 2022; Rong et al., 2023). For example, a lack of cultural sensitivity and competence from health care providers has been cited as a key factor leading to mistrust of the United States health care system by African Americans (Jacobs et al., 2006; Kennedy et al., 2007). Though effective community research also relies on mutual trust (Han et al., 2021), there is evidence that community-based participatory research can help to build trust and enable development of collaborative and equal partnerships between researchers and the community (Rong et al., 2023).

In addition to ensuring no bias of power, two-way information sharing might be considered a key characteristic of an equal partnership. There is some evidence that researchers commonly engage in research involving community engagement without providing feedback to their community partners (Mathie et al., 2018), despite the fact this is appreciated by partners and may motivate participants to engage further (Han et al., 2021). Dissemination of research findings to the community may be one way this could be achieved, with one study reporting that 73.2% of their sample of 109 community stakeholders reported "research results [which] are disseminated to the community in a culturally relevant and appropriate manner" as a top indicator of successful community engagement (Skinner et al., 2018). Building this information sharing into research might not only facilitate greater satisfaction of participants in community research but could lead to increased and broader engagement of other community members in research. As scientific findings are shared, trust is built, and knowledge and familiarity of research increases.

Another way to overcome systematic and structural biases may be to broaden the representation of researchers themselves; a strategy which relies on systematic and structural changes by institutes and organisations as well as smaller actions by individual researchers. Whilst there's evidence that introducing more PhD fellowships which offer financial stability may increase the racial/ ethnic diversity of both applicants and enrolees (Ecton et al., 2021), other work indicates that efforts to increase diversity among STEM researchers has been largely unsuccessful (Miriti, 2020). Such changes are difficult for an individual researcher to accomplish, but other recommendations to help overcome structural and systematic biases are more achievable. For example, one recommendation is to always report socio-demographic variables of participant samples included in research and to integrate these into study results and a discussion which emphasises limits to the generalisability of findings (Garcini et al., 2022). In addition, where research does investigate experience-related findings, it is important that researchers carefully consider their theoretical framework and approach these in a way which attempts to reduce stigmatisation of any particular group. Given evidence indicating SES-related differences in development and the plasticity of the brain to adjust to input (Hackman & Farah, 2009), it is perhaps likely that neurodevelopment may be conditional upon environmental, cultural and structural factors. Existing research commonly interprets SES differences as deficits or disruption to typical development, whereas there is evidence that such differences are the result of experience-based adaptations (Ellis et al., 2017). Considering environmental impacts on development from a strength-based, adaptive perspective could help to reduce biases in research and may lead to increased diversity of participation (Garcini et al., 2022).

5.1.3 Socioeconomic status and cognitive ability

A common approach in conceptualising environmental influences on development has been through consideration of SES, which is a measure of an individual's standing in society. A significant body of research has found evidence of a relation between SES and cognition in young children, with findings from large, as well as smaller, samples of children. For example, in a sample of over 10,000 children, researchers found evidence of differences in IQ between children from high and low SES backgrounds that were already apparent at 2-years-old (von Stumm & Plomin, 2015), whilst reading and mathematics ability differences were found between low, medium and high SES groups in a sample of over 20,000 kindergarteners (Duncan & Magnuson, 2012). Smaller studies have made similar findings, with evidence for a relation between SES and cognitive, motor and language ability in 15- to 30-month-olds (Wild et al., 2013) and between SES and language ability at 18- and 24-months-old (Fernald et al., 2013). This work has been reviewed in more detail in the introduction to this thesis (section 1.6.1), which discussed associations between SES and general cognitive ability as well as between SES and specific cognitive domains (language and executive functioning). There is thus considerable evidence that general cognitive ability may be susceptible to environmental influences and might therefore be a useful focus for support for children at risk of poor outcomes.

Furthermore, general cognitive ability may be particularly suited for remote assessment. Typical methods for assessing general cognitive ability involve either parent-reported questionnaires (i.e. in the Ages and Stages Questionnaire (Agarwal et al., 2020)) or researcher-led completion of a battery of short tasks designed to assess what a child can or cannot yet do (i.e. Mullen Scale of Early Learning (Mullen, 1995)). Whereas assessments of visual attention or executive functioning (EF) skills commonly use eye-tracking or reaction time measures which must be very accurate and precise, general cognitive measures gather a broader and more holistic view of cognitive ability. In this way, it may be possible to use parent-administered tasks and parent-reported performance on these tasks to effectively gather cognitive information about infants and young toddlers. It should be said that slightly older toddlers may be better able to use a touchscreen themselves. This means app-based tools hold additional potential for direct collection of data in tasks which gather measures of reaction time, and which may focus on specific domains such as EF. This chapter only includes data collected from an infant app that included tasks based around typical researcher-led assessments, however parents in focus groups were asked about a prospective toddler app which also included tasks to be completed by the child themselves.

Given considerable evidence of associations between SES and cognition, and the suitability of general cognitive ability for remote assessment, the current chapter explored associations between SES and cognitive ability assessed by an app-based tool for remote data collection. In light of the rapid brain development that occurs in the first years of life (see section 1.2.2), development of a tool that can efficiently and reliably measure cognition over this time may be useful for tracking rapidly changing abilities in a way that less-frequent lab-based studies cannot (Fearon, 2019). The potential for an environment-development link is also further supported by properties of the brain which indicate a potential sensitivity to environmental inputs during this period (see sections 1.5.1), yet many studies looking at SES-cognition relations have focussed on slightly older children than in the current chapter, therefore the current work extended knowledge to younger infants.

5.1.4 Current study

The current chapter focussed on the development of an effective, scalable tool for collecting developmental data from a diverse sample of infants and toddlers which suits both the needs of researchers and parents or childcare workers and can be used to explore the relation between early experiences and development. Specifically, it was focussed on developing an app-based tool suited for remote data collection in communities without a need to attend a lab or other setting. Such a tool could help overcome some of the challenges which may lead to some families not engaging in research and may ultimately lead to more diverse participation. Increased diversity in DCN research is needed to improve research such that it better represents the whole population, and to fully investigate the relations between early experiences and cognitive development.

This chapter comprised of three sub-goals which each contribute to the overall aim. The first focussed on understanding reasons why parents do or do not engage in research, and how future work could be best suited to support their participation. The second sub-goal aimed to assess how effective a current app-based tool was in collecting developmental data remotely from a diverse sample of infants. Findings from both these sub-goals might together inform about methods for improving diversity of representation in future developmental research. Finally, the third sub-goal took a slightly different focus and considered how SES features can be grouped together and whether these groupings may relate to any differences in profiles of cognitive ability. SES features in this study included objective measures of household income, people per bedroom, parental education, and occupation, as well as two subjective measures which were self-reported standing in both the local community and in the UK. Such an approach may inform about how elements of children's early experiences could impact their cognitive development.

This chapter included data from multiple studies which were combined to assess the three sub-goals. Data from a focus group study designed specifically to understand parents' views about research, child development and an existing version of an app designed for parents of toddlers aged between 18-months- and 3-years-old (named 'iTapp'), as well as questionnaire data collected during the lab-based study described in chapter 4, were merged to build an understanding about reasons why families might engage in research and how future work could be best suited to engage them. The latter two sub-aims were investigated using data from an online pilot study which involved an app (named 'Teachbrite') developed for parents of infants aged 4-18 months, as well as online questionnaires. Data from the Teachbrite app study were used to assess the effectiveness of an app-based tool in two main ways: (1) by assessing how effective the app was in reaching a wide range of participants and (2) by assessing the validity of the tasks in the app. Data were additionally investigated to explore associations between SES and cognitive ability.

In line with work suggesting that community engagement through the research process can lead to greater trust and may improve the diversity of participation (Rong et al., 2023), focus groups were conducted to build understanding about parental views of developmental research and to gather information about how an app-based tool could be optimally developed for use by researchers, parents and childcare workers in a variety of socioeconomic settings. It is hoped that incorporating community involvement in the development of this remote data collection tool will help the final app to be more user-friendly, which could lead to increased and more meaningful use. Furthermore, community involvement may help to establish and build meaningful relationships through which information can flow in both directions between researchers and communities. Participants of focus groups were specifically asked about what they would like to get from an app-based tool and how researchers could help share information to support them.

As part of the second sub-aim to assess how effective the Teachbrite app was in collecting developmental data remotely from a diverse sample of families, the sociodemographic profile of participants recruited to the study and of those who actively used the app was compared. This was done to provide to an understanding about how effective this remote tool might be in improving diversity in DCN research, and is in line with recommended practise to help reduce systematic biases towards certain communities (Garcini et al., 2022). Though a multitude of parenting apps exist, there remains a need to assess the usability, feasibility and acceptability of apps for all communities there are designed to serve (Crouse et al., 2023).

To assess validity of data collected by the Teachbrite app, associations between developmental data gathered from the app, data gathered by a developmental questionnaire and infant age were also explored. It was predicted strong positive associations would be found, indicating high validity. Measurement validity is an essential requirement of any data collection tool; checking this provided a better idea about the suitability of this tool as a scalable remote data collection tool. In addition, online questionnaire data from the Teachbrite app study were utilised to assess the relation between SES and cognition, a relation that has been broadly reported (Duncan et al., 1998; Duncan & Magnuson, 2012; Fernald et al., 2013). The current study used a cluster analysis to consider profiles of SES before using these cluster groupings to investigate whether there are any SES-related differences in cognitive profile. This analysis approach was motivated by a strength-based approach to environmental experiences, which has been recommended to help reduce systematic biases and improve accessibility of research to a wider range of families (Garcini et al., 2022).

In summary, the current chapter is concerned with the development of a scalable app-based tool which may be used in future research to increase diversity of participation in developmental cognitive neuroscience research and ultimately to better understand relations between early environmental experiences and development.

5.2 METHOD

5.2.1 Focus group study

5.2.1.1 Participants

Participants were 10 female parents recruited to take part in focus groups. Focus groups were conducted in a preschool in the UK. Early Childhood Education and Care settings were approached and asked if they would be willing to host a focus group; once a date and time was agreed, participants were recruited via information advertised both by the setting and the researcher. Though focus groups were held primarily for parents and workers associated with a setting, participants were not excluded if they were not associated with the setting. Focus groups were advertised to childcare workers and parents; whilst parents with children aged between 2-4 years were sought, parents of children up to 10 years old were included. Two focus groups were conducted, with 4-6 parents per focus group, totalling 10 participants in total (all female). Additional demographic information about the parent or family was not collected during the focus groups.

Formal signed consent was obtained from participants at the beginning of each focus group. Ethical approval was received for this study from the Department of Psychological Sciences ethics committee at Birkbeck, University of London (ref. no. 2223055_R).

5.2.1.2 Materials

5.2.1.2.1 Questionnaires

Three questionnaires were used; (a) one which contained questions about participants' views of research (Appendix D.1), (b) one which asked questions about participants' views of the factors influencing child development (Appendix D.2) and (c) one which related to the iTapp app and what parents would like from an app-based tool (Appendix D.3).

Among other questions, questionnaire (a) asked participants about the kinds of research they considered important. Developmental research was not specifically included as a type of research in this question, but related fields such as medical, genetic, and educational research were included, along with broader fields such as space and economic research. These were chosen as it was assumed that most participants would be interested in developmental research so these might provide insights about whether particular interests in related fields were driving participation in this study, and to build understanding about parents' wider view of research.

In questionnaire (b) parents were asked to rank the aspects of child development they were most interested in and to rank the factors that can impact children's development, in both cases from high (1) to low (11). The aspects of child development parents were asked to rank included the seven areas of learning and development that shape early years education as outlined by the Early Years Foundation Stage (EYFS) (Department for Education, 2021), as well as problem-solving, which is a key aspect of the EYFS' characteristics of effective learning. Three additional domains that are commonly considered by developmental researchers (attention, memory, and executive functioning) were also included. The factors impacting child development that parents were asked about were based upon common themes in developmental research and wider society.

Questionnaires were completed on pen and paper. Questionnaire data were inputted into an csv file and formatted ready for statistical analyses.

5.2.1.3 Procedure

Focus groups were conducted in a room of a preschool in the UK during summer 2023. Focus groups were planned to be an hour long, though this depended on participant contributions. Participants could leave at any time. Each session followed roughly the same format, with some flexibility depending on the direction which discussions took during the focus group. Parents were asked to complete questionnaires before discussions for each sub-topic such that participants could provide feedback anonymously and so that their answers were not unduly influenced by others' opinions during the discussion. The broad format was as follows:

- Introduction and welcome. Time for questions and consent. Recording started after consent.
- Questionnaires then discussion about views of research
- Questionnaires then discussion about the factors of child development that parents consider important

- Demonstration of the app and chance for participants to try the current version
- Questionnaire and discussion about the app
- Opportunity for questions and to take participants' contact information for future research if they agree. Thank you and goodbyes.

Recordings were taken of each session and, where a helper was available, notes were also taken in real time. These measures enabled the researcher to be focussed on the sessions and allow interactions to be more natural.

5.2.2 Lab-based study

Chapter 4 provides full methods for this study. In brief, participants were seventeen (seven female) 24- to 48-month-olds (mean age was 33.9 months) recruited for a study in the Birkbeck Toddlerlab. Here, the focus was on questionnaire data provided by parents of those toddlers, therefore consider parents to be the participants of the present study. The questionnaire included in this study was the parent feedback questionnaire, designed to assess parental views of research and child development, as well as the suitability of the study reported in chapter 4. Only questions relating to parental views of research and child development are included here; these were the same as those used in the focus group study a described in 5.2.1 above.

5.2.3 Teachbrite app study

5.2.3.1 Participants

Recruited participants were 124 parents recruited via an online website (https://psyc.bbk.ac.uk/teachbrite/evolutionstudy/index.php) which was hosted by Birkbeck University and could be freely accessed. In addition, this study was advertised on Birkbeck web and social media pages, as well as other webpages such as Children Helping Science (https://childrenhelpingscience.com/studies/). Recruitment information advertised the study to parents/ guardians of infants aged between 4 and 18 months. Parents/ guardians could access study information online and were required to provide informed consent before they could access study questionnaires and a link to download the app. Since the app was also freely available on the App Store and Google Play Store, a consent form was also built into the app to account for parents/ guardians who did not access the study website and downloaded the app directly. Only participants who were signed up via the website are included in this chapter. Parents were asked to provide age information only once they had accessed the app, therefore this information is only available for participants who used the app and not for participants who completed online questionnaires but did not use the Teachbrite app; demographic information is provided in section 5.3.2.1.

Ethical approval for this study was received from the Department of Psychological Sciences ethics committee at Birkbeck, University of London (ref. No 181962).

5.2.3.2 Materials and stimuli

5.2.3.2.1 Questionnaires

5.2.3.2.1.1 Sociodemographic questionnaire

There were two parts to this questionnaire (Appendix C.4); the first asked parents/ guardians to report on objective measures of SES, whilst the second asked for subjective responses about parent/ guardians' standing in life. Objective measures included questions about commonly used measures of SES such as income, education, employment and living circumstances. Subjective measures asked parents/ guardians to compare themselves to others (1) in the UK and (2) in their community by placing themselves on a scale between most well off and least well off. Participants were shown a visual representation of a ladder numbered from one to nine, with a score of one indicating those who are most well off and nine indicating those who are least well off.

5.2.3.2.1.2 Ages and Stages questionnaire (ASQ-3)

The ASQ-3 is a parent-completed questionnaire commonly used as a screening tool for general development. It is designed to assess the following five developmental domains: communication, gross motor, fine motor, problem-solving and personal-social.

The ASQ-3 is designed to cover children ages from 1 month to 5.5 years with 21 time intervals between these ages. For each interval there is an age-specific questionnaire containing a total of 30 items (six per domain). As the age range of the current study was relatively large, the questionnaire was collapsed across intervals at 2,4, 6, 8, 9, 10, 12, 14, 16, 18, 20, 22, 24 months. This meant questions were taken from questionnaires at each of these intervals and created one questionnaire with five domains (and one 'overall' section) which covered all ages (4-18 months) included in this study. As there is some crossover between questions from questionnaires at different intervals, the number of items in each domain ranged between 28-32 (28, 32, 32, 30, 28). To avoid parents/ guardians of younger infants answering questions relating to behaviours far beyond their child's current development, they were instructed to skip to the next domain once they had answered not yet' to four consecutive questions. The ASQ-3 typically takes around 10-15 minutes to complete: due to us collapsing across intervals, it was a little lengthier than normal however took somewhere between 15-25 minutes to complete. In each domain, the vast majority of items could be answered with a 'yes', 'sometimes' or 'not yet', though there were also a small number of open-ended items in some domains.

5.2.3.2.2 Teachbrite app

The Teachbrite app ('Teachbrite') was developed and built in Unity and uploaded to both the Apple Store and Google Play Store. When initially accessing the app, participants had to complete a built-in consent form, before entering the 5-digit ID number they were provided with during study registration (their unique ID code). The Teachbrite app asked participants for some basic personal information before proceeding to the 'task' section of the app.

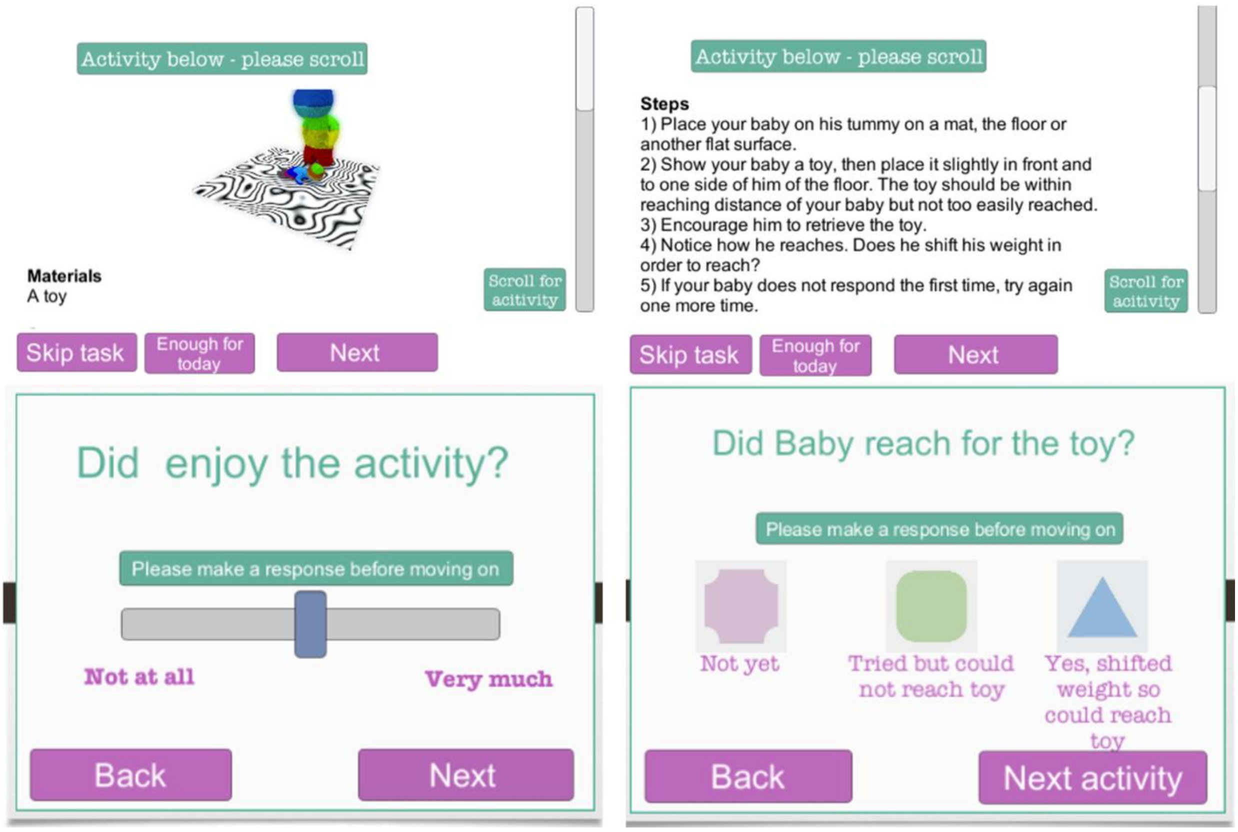
The app was initially developed as part of another project, but (besides piloting) this study was the first time this version of the app was used to gather data. Initial work on the app was in partnership with Babybrains (https://www.babybrains.info/), a company which specialises in communicating scientific advances to parents. The aim of the overall project was to develop a method of personalised communication to individual families, by individualising app content to be relevant to individual children. The app would gather parents' answers about their child's responses to suggested tasks; the goal was to individualise the app such that these responses would be used to select the next activity the app would suggest and to provide neuroscience-based information about a related aspect of development. The current study was a first step towards this in which the Teachbrite app contained tasks and parent responses about their child's performance but did not yet contain personalisation algorithms for task selection or science-based feedback.

Initial app development was done by Professor Rob Leech and then me, with tasks written by Dr Silvia Dalvit Menabe (BabyBrains), Professor Emily Jones and myself. Tasks were based around standardised cognitive assessments such as the Mullen Scales of Early Learning (Mullen, 1995), the Vineland Adaptive Behaviour Scales (Sparrow & Cicchetti, 1989) and the IBQ-R (Gartstein & Rothbart, 2003; Putnam et al., 2014) and were written in parent-friendly language and adapted to be easily executed with minimal resource requirements. These assessments are frequently used with infants and young children and cover a wide range of developmental domains (i.e. language and communication, temperament, socialisation), therefore they were well-suited to gather a broad assessment of infants' abilities. They additionally rely on a relatively small set of items to perform tasks, therefore could be easily adapted to use in homes with items that may be readily available. Tasks were stored in a csv-file which was read by Unity and used to fill in sections of the user interface, including resources required, instructions to parents and the image accompanying each task. The row number for each task in the csv-file was used as an index with which task responses were stored and tasks were randomly suggested by the app.

For each task, participants were shown sections containing resources required (or suggestions for these) and step-by-step instructions for completing the task, which were displayed with a cartoon-style image demonstrating the task (or similar). Images were developed by Professor Rob Leech and designed to show a generic parent and child. On each task page, participants had the option to proceed to the next page, go back a page or skip the task. If they chose to skip the task, the display changed to show another task; if they chose to proceed, the first of two question pages were displayed.

The first question page asked participants how much their infant enjoyed the activity, with a response bar showing a scale from 'not at all' to 'very much'. Following this, the second question page asked a task-specific question with a 3-point Likert response scale response. For most tasks, the 3 options largely mapped to responses indicating that a baby (1) couldn't yet do the task, (2) made some attempt but was not successful or (3) successfully completed the task, however this varied for some tasks to best fit the development each task was investigating. This 3-point response scale was based upon the scoring system for the Vineland Adaptive Behaviour Scales and could be easily adapted to others tasks in the app. Whilst many items in the Mullen Scales of Early Learning use only 2-point scale (whereby an infant scores zero if they cannot do the task or one if they can), some use a 3-point scale with an intermediate level of completion. As this intermediate level may be informative about how children are developing (i.e. it might indicate children are close to grasping a new skill but not quite yet), it was decided to use a 3-point scale in the Teachbrite app. It was felt the addition of additional levels (i.e. using a 5-point scale) was not justified based upon existing measures and may have caused uncertainty or confusion for parents. Figure 5.1 shows screenshots of the activity and questions pages of the Teachbrite app.

Figure 5.1: Example activity and question pages from a task in the Teachbrite app



When participants wanted to leave the app, they could choose to exit the app by pressing the ‘enough for today’ button and were given an option to confirm this on the following page. Participants could re-access the app at any given time and continue completing tasks, without needing to re-enter their personal information.

The Teachbrite app was designed to randomly suggest activities such that it could build a space about how performance on different tasks is related, with the aim to build an algorithm which would personalise later versions of the app. A large amount of data collected across tasks and stages of development is needed to achieve this aim, therefore the current study utilised a different approach to attain a cognitive score from the app to compare to questionnaire measures.

5.2.3.3 Procedure

Parents could access information about the study at a study website hosted by Birkbeck Psychology IT services: https://psyc.bbk.ac.uk/teachbrite/evolutionstudy/. Parents who chose to take part in the study were required to complete online consent before registering for the study. For study registration participants were asked to provide an email address. Email addresses were not associated with any questionnaire answers or app data but were saved in a separate list to be used to send reminder emails about the study. After submitting an email address, participants were given a randomly generated five-digit ID number which was both shown on screen and automatically emailed to the provided email address. Participants then proceeded to online questionnaires and were asked to complete the sociodemographic questionnaire followed by the ASQ-3. After completing questionnaires, participants were provided with links to download the Teachbrite app from the Apple Store or Google Play Store. Participants were asked to use the app as often as they could within the next two weeks, though they were able to continue using it after this point.

Personal information asked by the Teachbrite app included parent/ guardian name, infant name, age in weeks and gender, though names were only stored with the app on each device and were not saved as part of the study. Age in weeks and gender information were encrypted and sent to a secured server at Birkbeck where they were saved with their unique 5-digit ID code. Data from activities in the app were also saved with a participant's ID code in an area of the same server.

5.2.3.4 Data processing

5.2.3.4.1 Questionnaires

Data were downloaded from the Birkbeck server and incomplete or ingenuine data were removed. Ingenuine data were identified as files in which responses contained text that was irrelevant to the questions asked.

5.2.3.4.1.1 Sociodemographic questionnaire

Numeric data for household income, employment and education data were coded according to categories in the questionnaire. A metric of people per bedroom was calculated by dividing the number of people in the home by the number of bedrooms in the home. Further metrics of household income, education level and employment level were additionally determined. Household income data were divided into two groups relating to ‘high’ and ‘low’ income depending on whether their income was above or below £39,999. This split was chosen as roughly in line with the median household income in the UK in the 2022 (£38,100) (Office for National Statistics, 2023a). Education was split into two groups according to whether parents had or had not gone on to higher education, which is a category commonly used by the UK Office for National Statistics (Office for National Statistics, 2023b). Parent’s free text answers to what role they do were converted into occupation rankings based on (Hollingshead, 2011) occupational scale which ranks occupations from 1 (i.e. ‘farm labourers’, ‘service workers’) to 9 (i.e. ‘higher executives’, ‘major professionals’).

5.2.3.4.1.2 Ages and Stages questionnaire (ASQ-3)

In line with the standard ASQ-3 scoring, for questions which were not open-ended, items answered with ‘yes’ were scored with 10 points, ‘sometimes’ with 5 points and ‘not yet’ with 0 points. These points were totalled for each of the five domains: communication, gross motor, fine motor, problem solving and personal-social. Subscale scores could range from 280 (28 items with maximum score of 10) to 320 (32 items with maximum score of 10), depending on the sub-domain. Usually, domain scores would be compared to a specified age cut-off to determine whether a child passed or failed that domain. Due to collapsing across age intervals, subscale scores were simply the summation of scores on items in that domain. Overall ASQ score was the total of scores across the five subscales; the highest possible score was 1500 (280+320+320+300+280).

5.2.3.4.2 Teachbrite app

Responses for app tasks were saved with a task index which related to row number in the csv file of tasks. Task indices were used to find task rankings, with app scores calculated based on these task rankings. Though the ultimate goal of the Teachbrite app is to individualise content such that tasks are suggested based upon performance on previous tasks, implementing this requires a large amount of information upon which algorithms can be built and the current version of the app instead randomly suggested activities for families to try. As such, tasks were not completed in developmental order, with tasks becoming increasingly difficult until a child's ceiling level was reached, meaning a score akin to those typically calculated on developmental assessments could not be gathered. Additionally, personalisation was not yet implemented meaning that information about which tasks were suggested was not informative about children's ability. A method of measurement whereby tasks were ranked by difficulty (which equates to the developmental order in which they are typically achieved as in traditional assessments) was therefore developed, and infants were given an app score relating to the most difficult task which they successfully completed. In this way, app scores were thought to reflect an infant's highest ability level on the tasks they completed. Two app score measures were calculated based on two systems of task rankings. Measure 1 was calculated from Mullen Scales-based task rankings, whilst measure 2 was based on individual task rankings. Since tasks based on the Mullen Scales of Early Learning (Mullen, 1995) formed the largest proportion of tasks in the app, task number in the Mullen Scales was chosen as the basis for the first ranking system. Tasks based upon other measures (i.e. the Vineland Adaptive Behaviour Scales and the IBQ-R) were fitted into categories that were based upon the Mullen Scales according to their similarity to Mullen tasks. In case this system was oversimplified and missed out on subtle differences between tasks, measure 2 used task rankings in which each individual task was ranked. Whilst these were based loosely upon difficulty levels determined by the questionnaires upon which tasks were based, these also relied upon subject-specific knowledge by the researcher determining the rankings. Scores for both measures were the task ranking for the highest-ranked task which participants successfully completed. The number of tasks completed by each participant was determined by counting any task for which participants had provided a valid answer to the second question in the app.

5.2.3.5 Statistical analysis

5.2.3.5.1 Aim one: understanding parental engagement in research

Data from the focus group and lab-based studies were used to investigate reasons for parents' engagement in research and how development of an app-based tool might be most effective in encouraging participation. This focussed on three areas; (1) parental views of research, (2) parental views of child development and (3) parental views of an app-based tool; and involved both questionnaire data and discussion data from focus groups. For questions relating to parental views of research and child development, questionnaire data from the focus group and lab-based studies were analysed to compare responses from parents in the two groups. Most questions used rankings or ordinal scales, therefore Kruskal Wallis tests were used to compare groups. Questions relating to the app were only answered by participants in the focus group study, therefore descriptive statistics are presented for these. Discussion data from focus groups were analysed using a thematic analysis following the standard six-step process: (1) familiarisation, (2) coding, (3) generating themes, (4) reviewing themes, (5) defining and naming themes, (6) writing up (Braun & Clarke, 2006).

5.2.3.5.2 Aim two: assessing effectiveness of an app-based tool

To investigate the second aim of this study, to assess the effectiveness of the Teachbrite app in collecting developmental data from a diverse sample of children, data from the Teachbrite App study were used. Analyses first compared the demographic profiles of participants who signed up to the study and either did or did not use the Teachbrite app. The 'app users sample' included any participants who used the app regardless of whether they provided valid task data, whilst the 'not app users sample' was all other participants who provided online questionnaire data. Assumptions for tests were checked and appropriate independent comparison tests were chosen. For variable of people per bedroom, self-rated rankings in the community and in the UK, and ASQ score, data were not normally distributed, therefore Mann Whitney tests were used to compare between groups. For nominal variables of employments status, household income and education level, chi-square tests were used. Where contingency tables have fewer than five per condition, fisher's test is recommended instead of chi-square test (Kim, 2017); in the current study some conditions had exactly five participants, therefore fisher's and chi-square tests were both reported. Descriptive statistics for the whole recruited sample (whether they did or did not use the app) were also included.

To investigate the effectiveness of the Teachbrite app as a data collection tool for infants and young toddlers, data from the app were considered in relation to ASQ scores and age. The relation between ASQ scores and age was additionally investigated, to ensure collecting ASQ data in this way provided the expected pattern. The 'app task sample' included participants who provided valid data for any numbers of tasks; it differs slightly from the app users sample as 3 participants did use the app but did not provide usable task data. As a number of participants only provided a small amount of usable task data, a 'focussed app sample' was also determined as participants who provided data for at least 3 tasks. Relations between app scores and each of ASQ scores and age were investigated for both samples, with partial correlations additionally conducted to consider the relation between app and ASQ scores whilst controlling for number of tasks. Pearson correlations were used to analyse relations if all covariables were normally distributed, otherwise Spearman correlation was used. Analyses were repeated for both app score measures, therefore Bonferroni corrections were made for multiple analyses and significance was assessed at the $p < .004$ level.

5.2.3.5.3 Aim three: relation between SES and cognitive profiles

To investigate how SES measures were grouped, a hierarchical cluster analysis was conducted using measures of high or low household income, high or low education level, occupation ranking, people per bedroom and self-rated ranking in the community and the UK. Cluster groupings from this analysis were then used to investigate the relation between SES and cognitive abilities assessed by overall ASQ score and scores for subscales of communication, fine motor, gross motor, problem solving and personal-social.

5.3 RESULTS

5.3.1 Aim one: understanding parental engagement in research

The first focusses on understanding the reasons parents do or do not engage in research, and how future work could be best suited to support their participation.

5.3.1.1 Parental views of research

Most parents in both studies had some previous experience of research, with few parents having had either a little or a lot of experience and over 50% who had some experience of research post-school; Table 5.1. A Kruskal-Wallis test found no significant differences in the distributions of experience ratings between the two samples, $H(1) = 1.47$, $p = .226$, with mean ranks for the focus group equalling 15.70 and for the lab study group equalling 12.13.

Table 5.1: Frequency statistics for parents' previous experience of research

	Focus group sample		Lab study sample		Whole sample	
	N	%	*N*	%	*N*	%
None	0	0	1	6.3	1	3.8
A little (i.e. during school)	3	9.1	5	25.0	5	19.2
A little more (i.e. some at university or post school, but limited)	3	27.3	5	25.0	7	26.9
Some (i.e. conducted own research at university)	4	36.4	6	37.5	10	38.5
Lots (have conducted more than one study)	2	18.2	0	0	2	7.7
Extensive (have multiple years of research experience, etc.)	0	0	1	6.3	1	3.8

To understand parents' more general views of research and to analyse whether parents may have been influenced to take part in the study due to interest in a particular type of research, rankings about which kinds of research they found most important were considered; these revealed considerable individual variability. Kruskal Wallis tests found significant differences in group rankings for medical and space research only, with means and mean ranks indicating that medical research was rated as more important and space research was rated less important by the focus group compared to the lab study sample; Table 5.2, Table 5.3 and Table 5.4.

Table 5.2: Statistics from Kruskal Wallis tests comparing how the two samples differed in their ratings of types of research. Significance is indicated by a *

	H	*df*	*p*
Economic	0.77	1	.380
Medical	15.26	1	<.001*
Engineering (Robotics)	2.83	1	.092
Space	6.01	1	.014*
Chemistry	0.02	1	.881
Climate	3.05	1	.081
Education	3.34	1	.067
Genetics	1.72	1	.189
Sports and active living	2.59	1	.108

Table 5.3: Mean ranks for different types of research for the two sample groups

	Focus group sample mean rank	Lab study mean rank
Economic	13.95	11.46
Medical	6.00	17.14
Engineering (Robotics)	15.35	10.46
Space	16.55	9.61
Chemistry	12.75	12.32
Climate	14.63	9.71
Education	9.45	14.68
Genetics	10.30	14.07
Sports and active living	9.22	13.79

Table 5.4: Mean, standard deviation and sample size for rankings of different types of research for each of the two samples and the whole sample

	Focus group sample			Lab study sample			Whole sample		
	M	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*
Economic	5.2	2.3	10	4.4	2.6	14	4.7	2.5	24
Medical	1.3	0.7	10	5.3	2.2	14	3.6	2.6	24
Engineering (Robotics)	6.3	2.0	10	4.5	2.7	14	5.3	2.5	24
Space	8.0	1.7	10	5.6	2.9	14	6.6	2.7	24
Chemistry	6.5	1.3	10	6.2	2.0	14	6.3	1.7	24
Climate	4.6	2.3	8	2.9	2.3	14	3.6	2.4	22
Education	2.8	1.0	10	4.4	2.2	14	3.7	1.9	24
Genetics	3.9	1.4	10	5.0	2.1	14	4.5	1.9	24
Sports and active living	5.2	2.5	9	6.8	2.8	14	6.2	2.7	23

Across both focus groups, two themes relating to motivation for taking part became apparent; these were a specific personal interest in the project focus and previous experience of research in their own studies or work; Table 5.5.

Table 5.5: Focus group quotes for two themes relating to parent's motivation for taking part in the research study

PERSONAL INTEREST IN PROJECT	EXPERIENCE OF RESEARCH
…sounded like a really interesting idea	*Just finished my own undergraduate degree*
I think it could be something I would want to use	*Know about Montessori and worked as a nanny for a few years*
Want to make it better or easier for other parents	
Excited to see what it was about	*I work in marketing research*

5.3.1.2 Parental views about child development

Parent's level of interest in child development for the whole sample was high; M = 87.96, SD = 13.68, N = 26, min. = 40, max. = 100. A Mann-Whitney comparison test found no significant difference in scores between the focus group sample; M = 92.40, SD = 7.44, N = 10; and the lab study sample; M = 85.19, SD = 16.03, N = 16; U = 59.50, Z = -1.10, p = .286.

Parent rankings about areas of child development they were most interested in showed some differences between the two sample groups (Table 5.6 and Table 5.7). In particular, the focus group sample considered children's understanding of the world to be less important than the lab study group, whilst the lab study group considered attention to be less important than the focus group sample. Descriptive statistics indicate that emotional, physical and communication/ language development were generally ranked highly (i.e. more important) (Table 5.8). Of note, a number of participants found it too difficult to rank areas of development, citing that all were equally important; they have been left out of these analyses.

Table 5.6: Statistics from Kruskal Wallis tests comparing how the two samples differed in their ratings of different areas of child development. Significance is indicated with a *

	H	df	p
Emotional	0.70	1	.404
Physical	0.12	1	.735
Communication/ language	0.12	1	.726
Literacy	0.84	1	.361
Maths	1.27	1	.260
Creative/ imaginative	2.12	1	.146
Understanding the World	5.85	1	.016*
Problem Solving	0.93	1	.335
Attention	6.93	1	.008*

Memory	3.32	1	.069
Executive Functioning	3.27	1	.071

Table 5.7: Mean ranks for different areas of child development for the two groups

	Focus group sample mean rank	Lab study mean rank
Emotional	9.10	11.59
Physical	11.80	10.75
Communication/ language	11.80	10.75
Literacy	13.20	10.31
Maths	13.70	10.16
Creative/ imaginative	14.50	9.91
Understanding the World	16.75	8.94
Problem Solving	8.70	11.72
Attention	4.70	12.97
Memory	6.70	12.34
Executive Functioning	6.70	12.34

Table 5.8: Mean, standard deviation and sample size for rankings of different areas of child development for each of the two samples and the whole sample together

	Focus group sample			Lab study sample			Whole sample		
	M	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*
Emotional	2.00	1.73	5	3.06	2.57	16	2.81	2.40	21
Physical	4.40	3.58	5	3.38	2.78	16	3.62	2.92	21
Communication/ language	4.00	1.87	5	3.50	1.90	16	3.62	1.86	21
Literacy	7.40	2.70	5	6.25	2.91	16	6.52	2.84	21
Maths	8.00	2.74	5	6.25	2.75	16	6.67	2.78	21
Creative/ imaginative	8.40	3.78	5	6.69	2.68	16	7.10	2.97	21
Understanding the World	7.75	1.5	4	5.13	1.54	16	5.65	1.84	20
Problem Solving	5.20	1.79	5	6.63	2.83	16	6.29	2.65	21
Attention	4.60	1.52	5	8.25	2.30	16	7.38	2.64	21
Memory	7.20	2.78	5	9.19	2.29	16	8.71	2.49	21
Executive Functioning	4.80	3.35	5	7.81	3.33	16	7.10	3.51	21

Descriptive statistics relating to rankings about how interested parents were in different factors impacting development are in Table 5.10 and Table 5.11. Comparison tests found that parents in the lab study ranked the impact of sex/ gender and social/ cultural practices as more interesting than the focus group sample did; Table 5.9.

Table 5.9: Statistics from Kruskal Wallis tests comparing how the two samples differed in their interest in factors impacting child development. Significance is indicated by a *

	H	df	*p*
Sleep	1.33	1	.249
Diet/ Nutrition	1.37	1	.243
Genetics	0.00	1	1.00
Sex/ gender	4.66	1	.031
Exercise	0.39	1	.531
Parenting	2.45	1	.117
Social/ cultural practices	6.13	1	.013*
Education	0.08	1	.782
Family health/ stress	0.13	1	.723
Parental education, income, employment	1.81	1	.178
Siblings	0.40	1	.529

Table 5.10: Mean ranks for interest in factors impacting child development for the two sample groups

	Focus group sample mean rank	Lab study mean rank
Sleep	8.58	11.97
Diet/ Nutrition	8.58	11.97
Genetics	11.00	11.00
Sex/ gender	15.58	9.17
Exercise	9.67	11.53

Parenting	7.67	12.33
Social/ cultural practices	16.25	8.90
Education	10.42	11.23
Family health/ stress	10.25	11.30
Parental education, income, employment	8.17	12.13
Siblings	12.33	10.47

Table 5.11: Mean, standard deviation and sample size for rankings of interest in factors impacting child development for each of the two samples and the whole sample

	Focus group sample			Lab study sample			Whole sample		
	M	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*
Sleep	3.00	2.28	6	5.40	3.76	15	4.71	3.52	21
Diet/ Nutrition	2.83	2.14	6	3.87	2.75	15	3.57	2.58	21
Genetics	4.67	3.20	6	4.47	2.67	15	4.52	2.75	21
Sex/ gender	8.33	3.33	6	5.47	2.67	15	6.29	3.09	21
Exercise	5.83	2.79	6	6.80	3.34	15	6.52	3.16	21
Parenting	4.33	1.97	6	6.33	2.85	15	5.76	2.74	21
Social/ cultural practices	9.33	2.42	6	5.20	3.26	15	6.38	3.54	21
Education	5.17	2.48	6	5.80	2.34	15	5.62	2.33	21
Family health/ stress	6.33	2.25	6	6.67	2.97	15	6.57	2.73	21
Parental education, income, employment	7.33	2.73	6	8.53	2.70	15	8.19	2.70	21
Siblings	8.83	1.94	6	7.47	3.36	15	7.86	3.04	21

Two key themes relating to parental views about different aspects of child development arose from focus groups; these were whether developmental areas were specifically important for individual children or for specific ages; Table 5.12.

Table 5.12: Focus group quotes for two themes which arose about parent's view of different aspects of child development

INDIVIDUAL-SPECIFIC IMPORTANCE	AGE-SPECIFIC IMPORTANCE
I put genetics/ medicine [as most important] because one of my children has a medical issue so things that are important to you	*I've got an older child [too] … it shifts as they get older*
I think probably emotional development [as most important] because it's very difficult to manage my child's emotions	*I think education [is most important] and that's because my eldest is year one now*
I was looking at it completely from my point of view… this isn't going to be applicable to everyone. This is completely relating to my family	*Whatever age they are you're right in that moment*

5.3.1.3 Feedback about current version of the app

No parents reported that the current version of the app was difficult to use, with most reporting it was quite easy; Table 5.13 and all parents reported that information in the app was quite (*N* = 8, 80%) or very clear (*N* = 2, 20%). When asked about aspects of the app they would improve, parents gave many practical suggestions, with themes relating to appearance, functioning and type of activities; Table 5.14.

Table 5.13: Frequency data relating to how easy the app was to use

	N	%
Very difficult	0	0
Quite difficult	0	0
Neither easy nor difficult	1	16.7
Quite easy	4	66.7
Very easy	1	16.7

Table 5.14: Parents' comments about aspects of the app they would improve and how

quit button in games, more colours, not sure if it reads out instructions
development games i.e. jigsaw of Europe with names and colour coded, then take away colours to see if they remember
personalisation - based on child's interests and success scores/ results on previous games played. Wide variety of games
can be used with more than 1 sibling? Does the data gathered from the activities the children complete alone lead on to appropriate activities to complete together?
if child activities then led to suggested activities for their stage of development. It could assess if some areas of development behind compared to others, then suggest activities to promote development in those areas
across age ranges
perhaps more colourful/ eye catching to appeal to children when using it
design v function - reframe app to see user journey. Add stronger colours, perhaps more engaging for little ones
great!
how to exit part of the app. It could be musical or vocal to gain the child's attention

Parents all indicated that they would use the app with some frequency, though none said they would use it multiple times a day; Table 5.15. Two parents were not included in frequency statistics as they ticked multiple answers, with one indicating that regularity of app use would differ depending on their purpose for using the app. There was considerable variability in parent's rankings of factors that would impact their use of the app, as demonstrated by variance measures in Table 5.16. One parent suggested other factors to those already suggested in the questionnaires, indicating that more information and activities in the app would increase their likelihood of using the app.

Table 5.15: Frequency statistics about how often parents reported they would use the app

	N	%
Never	0	0
Rarely	0	0
Every few weeks	1	4.8
Once a month	1	4.8
Once a week	2	9.5
Multiple times a week	1	4.8
Once a day	1	4.8
Multiple times a day	0	0

Table 5.16: Descriptive statistics relating to parent's ranking of factors impacting their use of the app

	Min.	Max.	*M*	*SD*	Mean rank
Ease of use	1	7	2.25	2.05	3.29
Functioning of the app	1	7	2.86	2.61	3.50
Appearance	2	6	4.29	1.89	5.64
Quality of information	1	5	2.38	1.51	3.07
Quality of activities	1	5	2.63	1.30	3.50
Entertainment/ learning opportunities for child	1	7	3.75	2.12	4.36
Founded in science/ link to researchers	1	7	3.71	2.14	4.64

There was some variability in parent's rankings of reasons for using the app; Table 5.17. As a way of communicating with child developmental researchers, entertainment for their child and to gain information about their child's development compared to large numbers of other children were generally ranked higher than other reasons, whilst one parent additionally indicated that they'd like the app to suggest activities that promote their child's development. When asked about their primary goal for using this app, parents reported a range of reasons, with most stating something about promoting their child's development or learning; Table 5.18. Of note, only one parent said a primary motivation would be for reassurance as the parent, though this was a central theme during discussions.

Table 5.17: Descriptive statistics relating to parent's ranking reasons for using the app

	Min.	Max.	*M*	*SD*	Mean rank
Track child's development	1	4	1.57	1.13	1.93
Snippets of scientific information	1	4	2.29	1.25	2.64
To communicate with researchers	2	7	4.86	2.04	5.86
Links to further resources	1	5	2.71	1.70	3.07
Information about child's performance on activities	1	5	2.71	1.60	3.43
Entertainment for child	2	7	4.86	2.19	5.93
Comparison to other children/ norms	2	7	4.57	1.99	5.14

Table 5.18: Parents' comments about their primary reason for using the app

helping to structure learning, ideas for games
to learn but enjoying doing so
to learn and help my child develop
to find enjoyable activities to complete in a short time frame, quickly, easily, and not too messy
aid my child's development
to help my child's development
fun & learning for the child, reassurance as the parent
entertain little one, learning developmental milestones
both for fun and to learn
understanding progress of my child

During focus groups, two key themes arose relating to parents' views for the purpose and use of an app-based tool. The first theme was to provide support for parents during the first few years of a child's life as parents felt this was currently lacking, whilst the second theme was personalisation of a tool to a child's development: Table 5.19.

Table 5.19: Focus group quotes for two themes relating to parent's view for the purpose and use of an app-based tool

TO PROVIDE SUPPORT FOR PARENTS IN CHILD'S FIRST FEW YEARS	PERSONALISATION OF TOOL
When we are pregnant, we could go to NCT [antenatal] classes to talk about what will happen when you were pregnant and when you give birth… and then there was nothing for after like how to help with sleep, what they should be eating…	*If it could identify if there were areas of development where they were a little bit behind compared to others and suggest some activities to train*
I think you think am I doing it right and should she have done that? Whereas if you have something there that could reassure and for your own mental health… because it is overwhelming especially if you had a bad night sleep	*I think it has to be more relevant to your child…So you see their own progressions. So then let's say people been using the app for a year then you can see they've really progressed in this area or that stayed pretty stable…not competing against anybody but just within your child so you can track what my child's done last year.*
For the reassurance	*If you identify like strengths and weaknesses, like with the young ones like say they're doing very well with fine motor skills. And then what, like, if there was a way of finding what maybe to work on*
When they start school and to some extent preschool there's lots of help whereas before that it's kind of like oh, they're eating and getting taller so it's fine.	
I had quite a few things especially with my first child, I needed some input from somebody but there's no one really to go to for it.	*I'd be completely on board with letting them play with an app if it was going to say they're at this level, do these things and it will improve this.*
I think it's more having the support for when something's not typical and it's not necessarily the negative. My oldest did quite a lot of things earlier than anticipated… and I didn't know how to deal with this... and the health visitors look at you like you're insane	*Parents are quite biased in the activities they give their kids to do…I think it would be good to have more suggested ideas of other areas*

Another key theme that emerged during discussion of an app-based tools for parents to use was the need for it to be suitable for multiple children within the family; Table 5.20.

Table 5.20: Focus group quotes relating to the theme of the suitability of an app-based tool for multiple children

SUITABILITY FOR MULTIPLE CHILDREN
If you've got multiple children, like I've got a got big age range. It can be hard to think of an activity for them all
Whether there's a way of setting up so you can add in multiple children and then it can look across multiple children and give groups suggestions too.
There's quite an age range and quite an ability range
...also thinking about siblings as well. Or if it's only if you get one-on-one time with that child

In relation to information sharing, parents generally indicated that they'd be willing to share most of the personal information suggested and as commonly reported in developmental psychology research, with a couple preferring not to share income, race/ ethnicity and medical history data (Table 5.21). In focus groups a key theme was that the relevance and purpose of collecting information was important in deciding whether to share (Table 5.22).

Table 5.21: Number and percentage of parents who indicated they would be willing to share each type of personal data

	N	%
Parental education	9	100
House size	8	89
Family size	9	100
Employment	8	89
Income	7	78
Race/ ethnicity	7	78
Medical history	7	78
Resources in home	8	89

Table 5.22: Focus group quotes for relevance and purpose of information sharing in an app-based tool

RELEVANCE AND PURPOSE OF INFORMATION IMPORTANT
…as long as the information that is being requested is relevant and purposeful
If I knew the source of where it was going, I would be comfortable with that
As much data as is going to be helpful
I recently did [something] for my child…and the information in the questions they sent me just didn't feel relevant at all... I think it has to be relevant as well otherwise you disengage

5.3.2 Aim two: assessing the effectiveness of an app-based tool

The second sub-goal aims to assess how effective two current app-based tools were in collecting developmental data from a diverse sample of children. The socioeconomic demographic profiles of recruited samples were first considered, then the relations between app data and age, and between app data and cognitive ability were investigated.

5.3.2.1 SES demographic profile of Teachbrite study

Analyses were performed to investigate whether there were differences between participants who signed up to the Teachbrite study and did or did not then use the Teachbrite app. Variables of people per bedroom, rankings in the community and UK, and ASQ scores were not normally distributed for one or both groups, therefore Mann Whitney tests were used to compare between groups. Analyses found no significant differences between the two groups for each of these variables (Table 5.23 and Table 5.24).

Table 5.23: Results from Mann Whitney tests comparing demographic data between the sample of not app users and app users

	W	*p*	CIs	
People per bedroom	1472.0	.320	-7.19	2.50
Self-rated ranking in community	1306.5	.988	-1.00	1.00
Self-rated ranking in the UK	1308.0	.995	-1.00	1.00
ASQ score	1296.5	.665	-130.00	180.00

Table 5.24: Descriptive statistics for demographic data for the whole sample recruited for the Teachbrite study, for app users and for not app users

	Not app users					App users					Whole recruited sample				
	M	*SD*	Min.	Max.	*N*	*M*	*SD*	Min.	Max.	*N*	*M*	*SD*	Min.	Max.	*N*
People per bedrooms	1.40	0.60	0.5	4	97	1.26	0.50	1	3	27	1.47	0.59	0.5	4	124
Self-rated ranking in community	4.43	1.50	1	7	97	4.48	1.19	2	7	27	4.44	1.43	1	7	124
Self-rated ranking in the UK	4.34	1.68	1	9	97	4.30	1.73	1	7	27	4.33	1.69	1	9	124
ASQ score	704.0	393.0	75	2055	75	676.0	358.0	225	1450	27	697.79	384.32	75	2055	118

Any differences between groups for nominal variables employment, household income and education were investigated using Chi-square and Fisher's tests (see section 5.2.3.5). Tests revealed that the proportions of participants with high and low income did not differ between those who did and did not use the app; nor did the proportions of participants who had or hadn't received higher education, or who were or were not currently employed (Table 5.25 and Table 5.26). Whilst participants were categorised in this way to enable these comparisons, distributions across broader categories are shown in Table 5.27, Table 5.28 and Table 5.29.

Table 5.25: Results from chi-square and fishers test comparing employment status, household income and education between app users and not app users

	Chi-square			Fisher's			
	χ^2	df	*p*	Odds ratio	*p*	CIs	
Employment status	0.26	1	.609	1.52	.610	0.49	5.71
Household income	0.41	1	.523	0.57	.339	0.16	2.34
Education	0.26	1	.614	0.62	.526	0.18	2.50

Table 5.26: Median and interquartile range for demographic data for the whole sample recruited for the Teachbrite study, for app users and for not app users

	Not app users		App users	
	Median	IQR	Median	IQR
People per bedrooms	1.33	0.5	1.33	0.73
Self-rated ranking in community	5	2	4	1
Self-rated ranking in the UK	4	2	5	2
ASQ score	115	62.5	95	70

Table 5.27: Frequency statistics for employment status of the whole sample recruited for the Teachbrite study, of app users and of not app users

	Not app users		App users		Whole recruited sample	
	N	%	*N*	%	*N*	%
Disabled (not working due to permanent or temporary disability)	3	3.1	1	3.7	4	3.2
Sick leave	0	0	0	0	0	0
Homemaker	17	17.5	3	11.1	20	16.1
Retired	0	0	0	0	0	0
Not currently employed, looking for work	5	5.2	1	3.7	6	4.8
Working part time	20	20.6	11	40.7	31	25.0
Working full time	49	50.5	10	37.0	59	47.6
Do not wish to answer	3	3.1	1	3.1	4	3.2

Table 5.28: Frequency statistics for household income of the whole sample recruited for the Teachbrite study, of app users and of not app users

	Not app users		App users		Whole recruited sample	
	N	%	*N*	%	*N*	%
<£20,00	1	1.0	4	14.8	5	4.0
£20,000 - £29,999	3	3.1	0	0	3	2.4
£30,000 - £39,999	7	7.2	1	3.7	8	6.5
£40,000 - £59,999	18	18.6	2	7.4	20	16.1
£60,000 - £79,999	9	9.3	3	11.1	12	9.7
£80,000 - £99,999	14	14.4	4	14.8	18	14.5

£100,000 - £149,999	22	22.7	6	22.2	28	22.6
> £149,999	11	11.3	4	14.8	15	12.1
Don't know or do not wish to answer	12	12.4	3	11.1	15	12.1

Table 5.29: Frequency statistics for education level of the whole sample recruited for the Teachbrite study, of app users and of not app users

	Not app users		App users		Whole recruited sample	
	N	%	*N*	%	*N*	%
Didn't finish secondary school	1	1.0	0	0	1	0.8
Didn't finish secondary school, but completed a technical/ vocational program	1	1.0	0	0	1	0.8
Secondary school – finished school but fewer than 4 passes in GCSEs, O levels or equivalent	0	0	0	0	0	0
Secondary school – passed at least 4 GCSEs, O levels or equivalent	4	4.1	4	14.8	8	6.5
Further education – completed A-levels, BTECs or equivalent	5	5.2	1	3.7	6	4.8
Higher education – undergraduate level (undergraduate degree, diploma, NVQ, etc)	26	26.8	8	29.6	34	27.4
Higher education – postgraduate level (masters, doctorate, PGCE, etc)	59	60.8	14	51.9	73	58.9
Do not wish to answer	1	1.0	0	0	1	0.8

5.3.2.2 Relation between app-based measures, ASQ scores and age

To assess the feasibility of collecting cognitive data via the Teachbrite app, analyses were done to compare app-based scores with age and ASQ score. Two different app-based measures were used; these are both based on ranking of task difficulty, with two systems of difficulty rating used. Task rankings and subsequent measures were (as expected) closely related; $r(119) = 0.99$, $p < .0001$ and $r(22) = 0.9$, $p <$

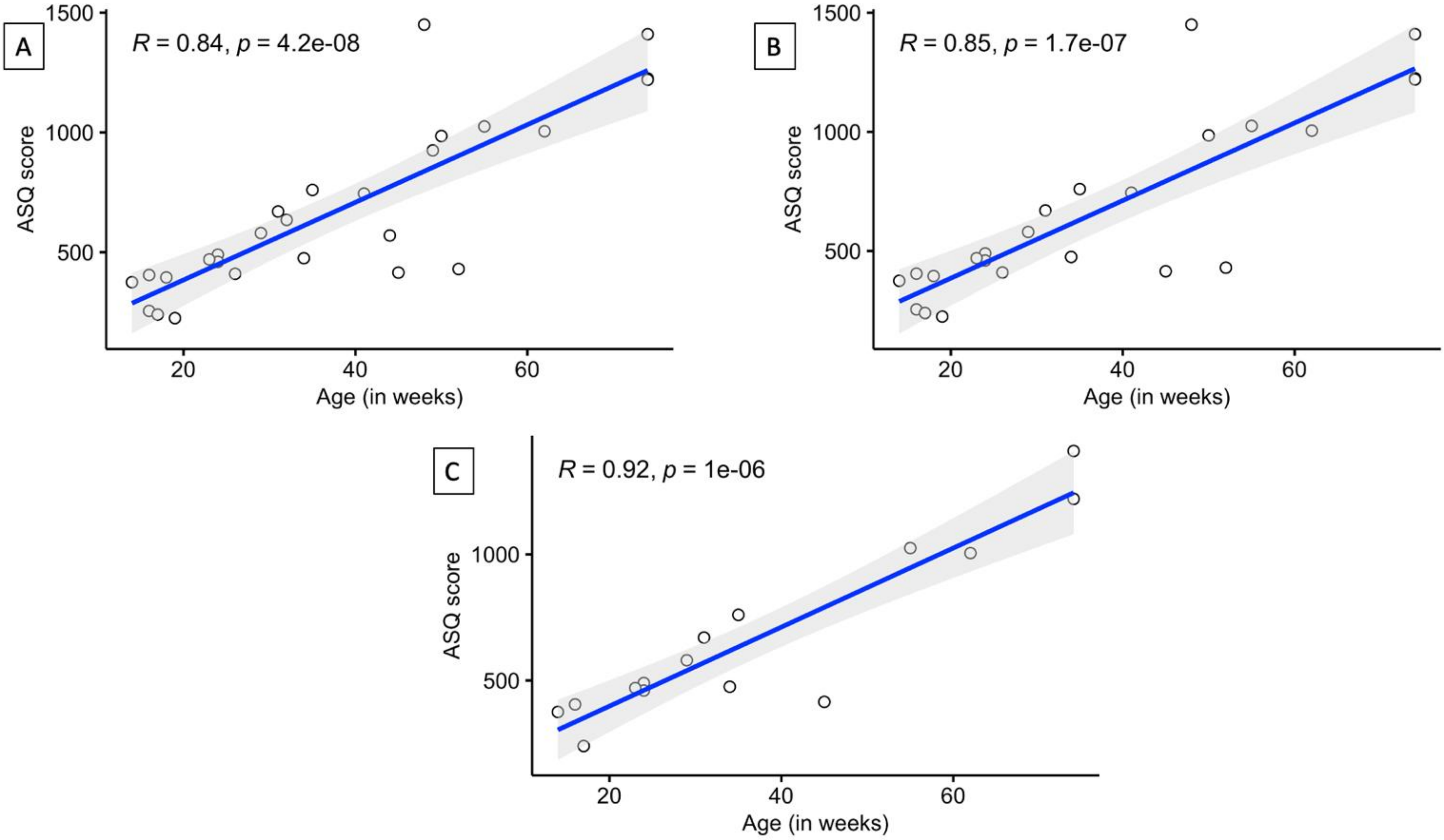
.0001; and both found similar results with age and ASQ score, as reported in the sections below. Descriptive statistics for related measures are reported in Table 5.30.

Table 5.30: Descriptive statistics for age, ASQ score, app scores and number of app tasks completed by each of the app task and focussed app samples

	App task sample					Focussed app sample				
	M	*SD*	Min.	Max.	*N*	*M*	*SD*	Min.	Max.	*N*
Age (in weeks)	37.54	19.7	14	74	24	37.13	20.3	14	74	15
ASQ Score	671.67	376.4	225	1450	24	666.67	345.0	240	1410	15
App score: measure 1	56.88	29.8	7	109	24	58.40	31.4	12	109	15
App score: measure 2	7.42	3.5	1	14	24	7.80	3.7	2	14	15
Number of app tasks completed	11.29	11.8	1	37	24	17.07	11.5	3	37	15

Across the whole sample who signed up for the Teachbrite study online, a Spearman's correlation found a significant positive relation between age in weeks and ASQ score; $r(25) = .84$, $p < .0001$: Figure 5.2a. A similar significant association was also found for the group who provided any app data; $r(22) = .85$, $p < .0001$; Figure 5.2B; and for the focussed app group (Pearson); $r(13) = .92$, $p < .0001$; Figure 5.2C.

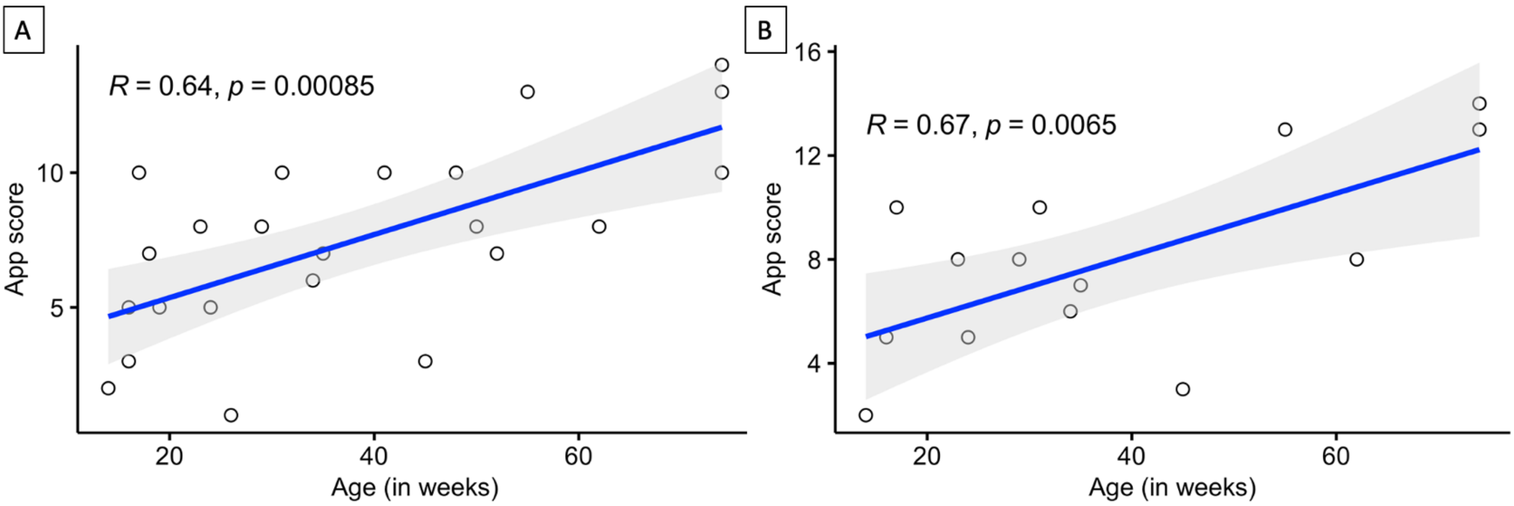
Figure 5.2: Correlation between age and ASQ score for (A) app users sample, (B) app task sample and (C) focussed app sample



5.3.2.2.1 App measure 1: Mullen-based ranking

A spearman correlation found a significant positive association between app score and age in weeks in the app study sample; $r(22) = .64$, $p = .001$; Figure 5.3A; Pearson correlation found a similar relation in the focussed app sample, though this was not significant when considered against the corrected significance level of $p < .004$; $r(13) = .67$, $p = .007$; Figure 5.3B.

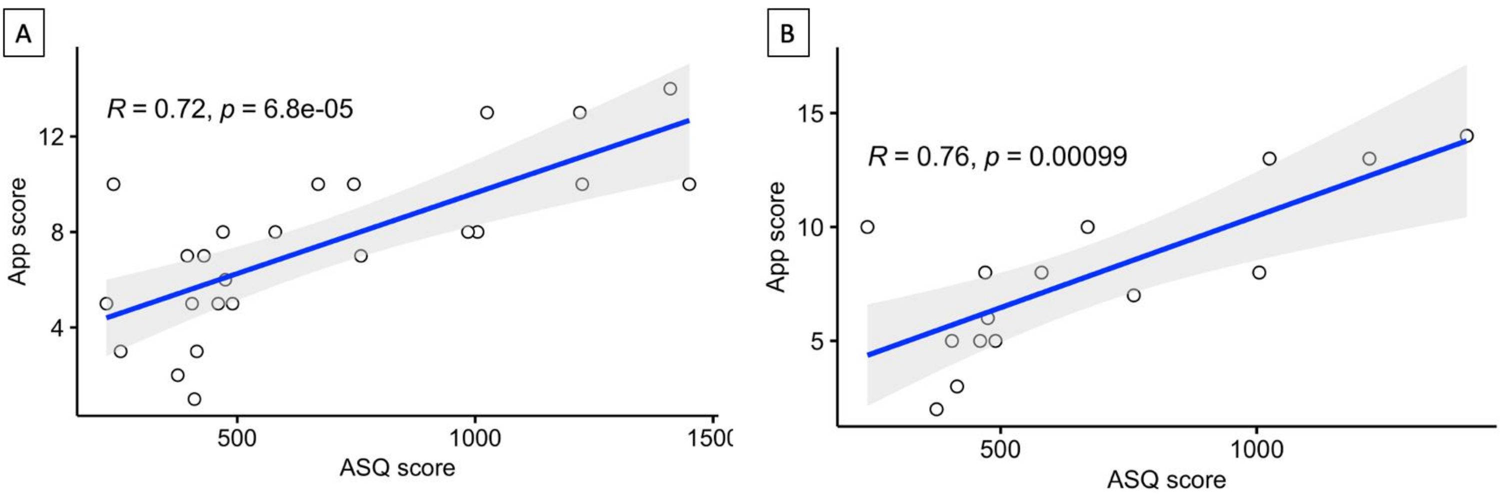
Figure 5.3: Correlation between age and Mullen-based app score for (A) app task



sample and (B) focussed app sample

A spearman correlation found a significant positive association between app score and ASQ score in the app study sample; $r(22) = .72$, $p < .001$; Figure 5.4A; a Pearson correlation found a similar relation in the focussed app sample; $r(13) = .76$, $p = .001$; Figure 5.4B. To further assess whether the number of tasks a participant completed impacted the relation between ASQ score and app score, a partial correlation was conducted. This found that a significant relation between ASQ and app score remained when number of tasks was controlled for; $r(21) = .69$, $p < .001$.

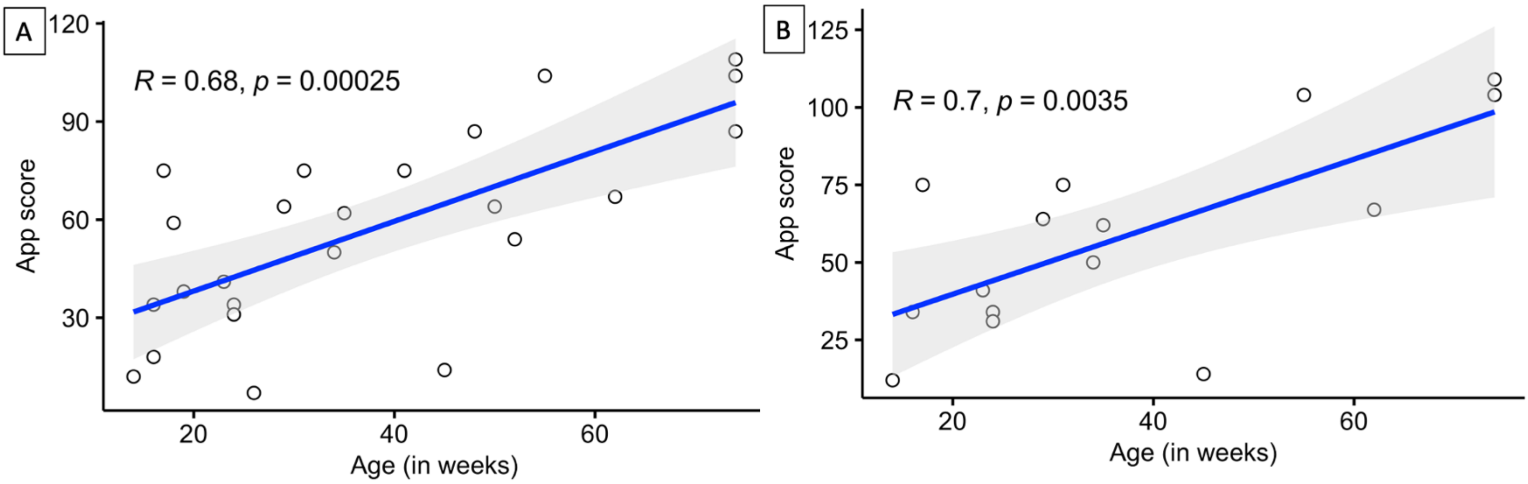
Figure 5.4: Correlation between ASQ score and Mullen-based app score for (A) app task sample and (B) focussed app sample



5.3.2.2.2 App measure 1: individual task ranking

A spearman correlation found a significant positive association between app score and age in weeks in the app study sample; $r(22) = .68$, $p < .001$; Figure 5.5A; a Pearson correlation found a similar relation in the focussed app sample, though this is very close to the corrected significance level of $p < .004$; $r(13) = .70$, $p = .004$; Figure 5.5B.

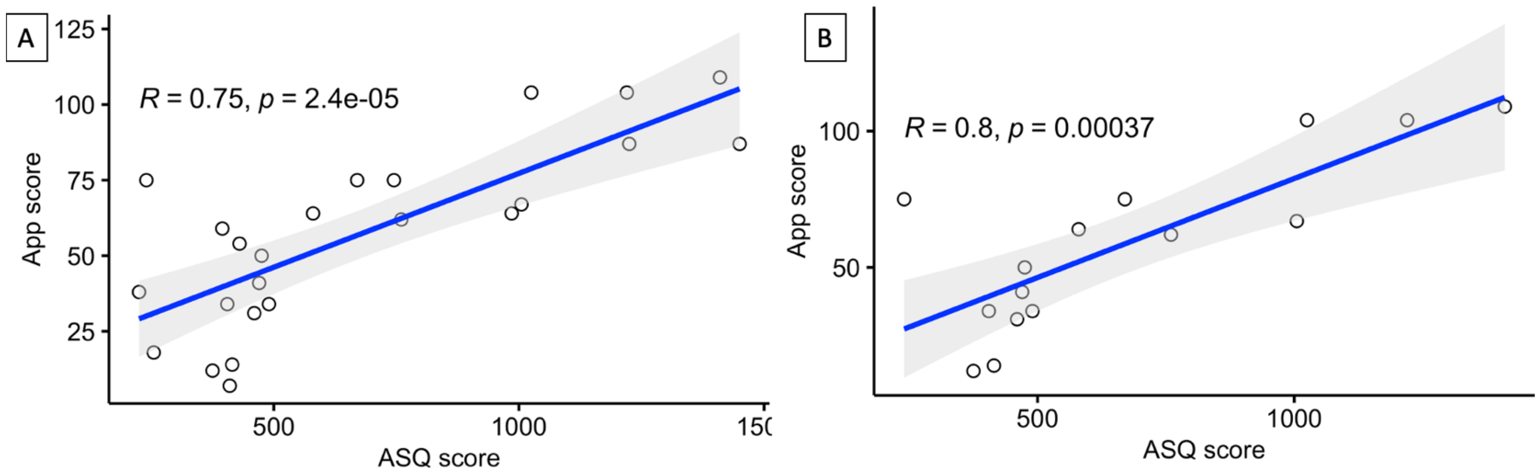
Figure 5.5: Correlation between age and individual task ranked app score for (A) app task sample and (B) focussed app sample



5.3.2.2.3 ASQ and app scores

A spearman correlation found a significant positive association between app score and ASQ score in the app study sample; $r(22) = .75$, $p < .001$; Figure 5.6A; a Pearson correlation found a similar relation in the focussed app sample; $r(13) = .79$, $p < .001$; Figure 5.6B. To further assess whether the number of tasks a participant completed impacted the relation between ASQ score and app score, a partial Spearman correlation was conducted. This found that a significant relation between ASQ and app score remained when number of tasks was controlled for; $r(21) = .72$ $p < .001$.

Figure 5.6: Correlation between ASQ score and individual task ranked app score for



(A) app task sample and (B) focussed app sample

5.3.3 Aim three: relation between SES and cognitive measures

Parents who did not provide any information about the nature of their work were excluded; any other responses which were empty, 'not given', 'NA' or '-999' were categorised as missing and missing data were imputed using the 'argImpute' function in RStudio (Harrell, 2023). Variables were scaled to have a mean of zero and a standard deviation of 1. A hierarchical cluster analysis was then conducted using the 'hclust' function in RStudio (Nowakowski, 2023). Household income, parental education level, occupation level, self-rated ranking in each of the UK and the community, and people per bedroom were used; see Table 5.31 for further information about each of these metrics. The Gower distance metric was calculated using the 'daisy' function in the cluster package in RStudio (Maechler et al., 2022) and a hierarchical cluster model was fit using the complete method. The optimal number of clusters was determined using the 'fviz_nblust' function in the factoextra package in RStudio (Kassambara & Mundt, 2020), which considers different methods including silhouette and gap statistics. This indicated two clusters were optimal. Cluster allocations were saved such that each participant was allocated to one of the two clusters. Table 5.31 shows unscaled mean values of each SES measure across the two clusters.

Table 5.31: Descriptive statistics for unscaled SES variables included in the cluster analysis

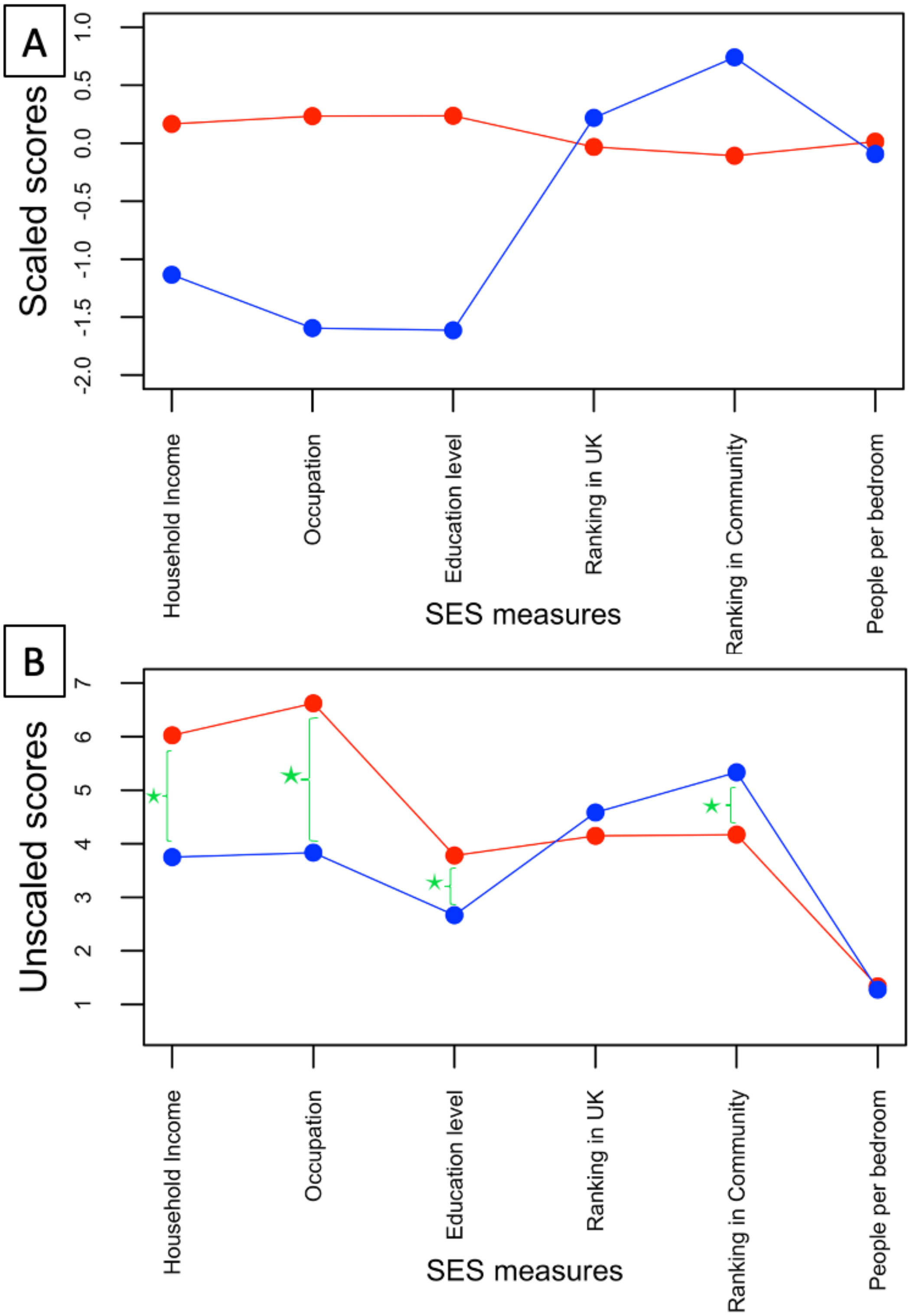
	Cluster group 1			Cluster group 2		
	M	*SD*	*N*	*M*	*SD*	*N*
Household income	6.02	1.57	82	3.75	1.66	12
Occupation level	6.60	1.12	82	3.83	1.75	12
Education level	3.78	0.42	82	2.67	0.78	12
Self-rated ranking in the UK	4.14	1.81	82	4.58	1.24	12
Self-rated ranking in the community	4.17	1.35	82	5.33	1.07	12
People per bedroom	1.34	0.62	82	1.28	0.33	12

Mann-Whitney tests compared unscaled scores for individual SES measures between cluster group one and two. These, together with descriptive data in Table 5.31, indicated that household income, occupation and education level were significantly higher in group one versus group two, whilst group two tended to self-rate themselves as of higher standing in the community than group one; Table 5.32. Figure 5.7 displays scaled and unscaled mean values for each SES measure for the two clusters, with stars indicating significant differences in unscaled scores.

Table 5.32: Statistics for Mann-Whitney tests comparing unscaled SES variables between cluster one and two

	W	*p*	CIs	
Household income	823	.0001	1.00	3.00
Occupation	903	<.0001	2.00	4.00
Education level	871	<.0001	1.00	1.00
Ladder UK ranking	384	.215	-2.00	2.72
Ladder community ranking	240	.003	-2.00	-3.04
People per bedroom	474	.841	-0.25	0.20

Figure 5.7: Mean scores for each SES variable in cluster groups one and two for (B) scaled scores and (B) unscaled scores. The red lines show scores for cluster group one, blue lines showing group two

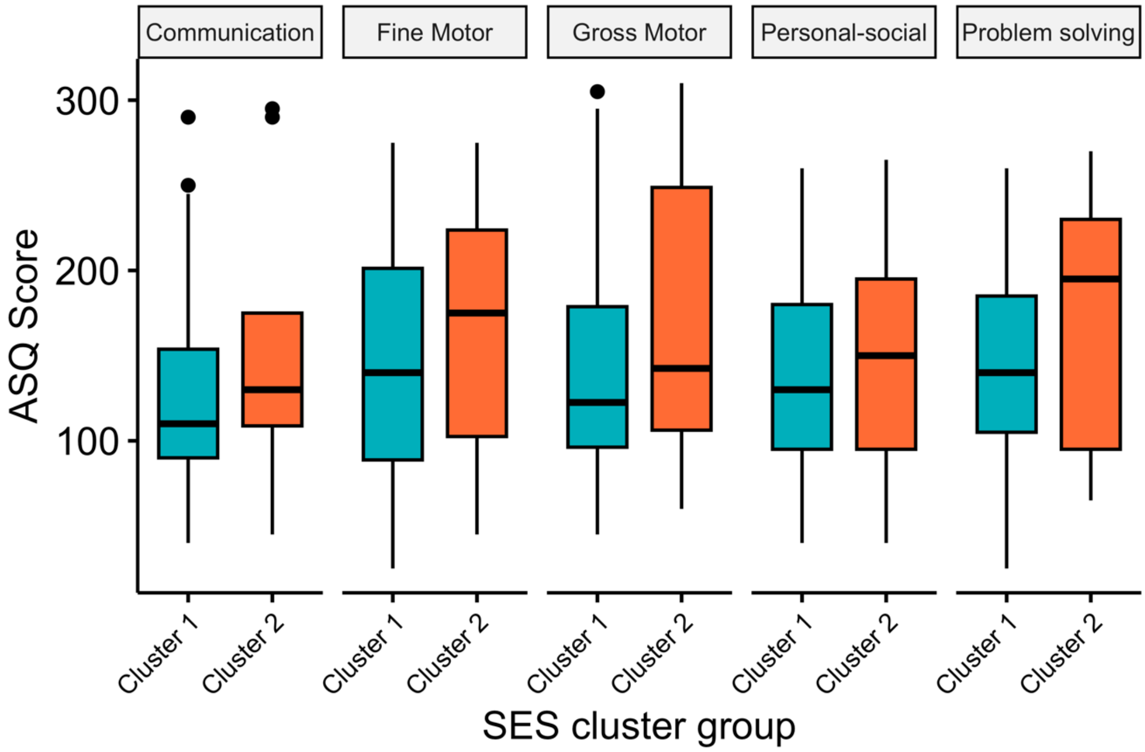


Analyses were also performed to compare cognitive scores between the two SES cluster groups. A Welch's t-test found no significant differences in overall ASQ score between the two groups: $t(81) = -0.71$, $p = .496$, $d = -0.32$. A multivariate ANOVA (MANOVA) was conducted to compare scores on ASQ subscales between the two SES cluster groups. Scores on the fine motor and problem solving subscales were highly correlated ($r = 0.92$), therefore problem solving scores were removed from the MANOVA in line with recommendations about avoiding multicollinearity (Tabachnick & Fidell, 2019). Communication, gross motor, fine motor, and personal-social subscale scores were dependent variables and SES cluster group was an independent variable with two levels (group one and two). The MANOVA found no significant differences in ASQ scores between the two SES groups; $F(4, 70) = 1.74$, $p = .152$; descriptive statistics are reported in Table 5.33 and displayed in Figure 5.8.

Table 5.33: Descriptive statistics for overall and each ASQ subscale for clusters one and two

	Cluster group 1			Cluster group 2		
	M	*SD*	*N*	*M*	*SD*	*N*
Overall ASQ score	680	280	66	818	403	9
Communication	125	54.1	78	151	75.9	12
Gross Motor	141	65.4	74	179	92.8	10
Fine Motor	143	66.8	72	162	85.0	10
Personal-social	137	58.3	66	144	74.8	9
Problem solving	144	56.1	70	166	76.5	9

Figure 5.8: Mean scores for cluster one and two for each ASQ subscale



5.4 DISCUSSION

This chapter focussed on the development of an effective, scalable, app-based tool for conducting developmental cognitive neuroscience research with a diverse sample of infants and toddlers which can be used to explore the relation between early experiences and development.

5.4.1 Aim one: understanding parental engagement in research

Focus groups and questionnaires from the focus group and lab-based studies revealed considerable variation in parental views of developmental research and how an app-based tool may be used in future research. Most parents had previous experience of research, though it was encouraging that one parent was taking part in research for the first time and others had only a little prior experience. Group comparisons found some differences between samples of parents from the two studies, with the focus group sample rating medical research more important and space research less important, children's understanding of the world less important and attention more important, and the impact of sex/ gender and social/ cultural practices, as less important than the lab study sample. These findings demonstrate potential differences in views between different communities which may be due to different experiences of parents in these areas. Though no sociodemographic data were collected from the focus group sample, focus groups took place in rural North Yorkshire and involved parents from the local community, whilst the lab-based study took part in the Birkbeck Toddlerlab in central London and involved parents from around London who had previously signed up to the Birkbeck Babylab and Toddlerlab database. The 2021 UK census found that 94.4% of the population in the local authority where focus groups took place were white, whilst this was true for 57.8% of the population in the local authority where the Birkbeck Toddlerlab is (Office for National Statistics, 2022). That participants in focus group study typically rated the impact of social/ cultural factors on development as less important than parents in the lab-based study may be reflective of reduced experience and less regular interaction with people from different cultures, due to the demographic profile of their local area. Such findings reveal differences in communities which emphasise a need to communicate with members of different communities to fully understand specifics of that community and demonstrate the importance of diverse representation in research.

Focus group discussions revealed two key themes underlying the areas of child development which parents considered most important, with relevance to individual differences between children, and age or stage of development at any given time shaping parental views. Following this, it is not surprising that personalisation of an app-based tool also emerged as a key theme in discussions about the purpose and uses of an app. Whilst parents were less clear about the details of personalisation, many reported a desire for an app which could identify their child’s strengths and weaknesses and track progress. In addition, despite individual and age differences impacting which areas of development parents considered important, there was considerable consensus indicating parents felt there was a lack of support and access to information during their child’s early years. This indicated a clear gap in parents’ needs which research may be able to help fill. Community partners have indicated that they appreciate feedback from researchers (Han et al., 2021) and consider it an important factor in successful research (Skinner et al., 2018), however researchers often fail to provide this (Mathie et al., 2018). Including built-in dissemination in a research tool may not only help to share useful research findings but could also build trust between researchers and communities (Rong et al., 2023). Increased trust within communities could lead to broader engagement, as well as potentially deepening the engagement of existing participants. This could play a crucial role in the development of equal and reciprocal partnerships between researchers and community members, which may ultimately lead to more representative research.

In relation to an existing version of a toddler app (iTapp), parents made numerous practical suggestions for improvements to appearance and appeal to young children, whilst conversations around the suitability of an app that can be used with multiple children in a family also emerged during focus groups. Whilst this may be practically more challenging to implement, it provides insight about key considerations families may have when engaging in research. Whilst a need for childcare was already apparent for families attending research in laboratory settings (Garcini et al., 2022), the current finding highlights that such practical difficulties may not be completely removed by using remote research methods.

Focus group parents indicated they would be comfortable sharing most types of personal information they were asked about, with all agreeing to share parental education and family size information and fewer agreeing to share income, race/ ethnicity, and medical information. These findings might reflect perspectives about the sensitivity of different types of information or could relate to parents' views about the relevance of information sharing, which was supported by focus group discussions around this topic. Many parents stated that their willingness to share data related to the purpose or relevance of doing so, with some indicating that they'd be happy to share any data as long it was made clear how it might help researchers. It is important to acknowledge that the types of data participants are willing to share may differ between across cultures and communities, which may in part be linked to historical experiences with research organisations and may be linked to concerns around misuse of information or stigmatisation (Garcini et al., 2022). Whilst the present findings are helpful in highlighting the need for researchers to clearly communicate the purposes for each element of their study, they should not be assumed to be true of communities other than those sampled in the current study.

5.4.2 Aim two: assessing the effectiveness of an app-based tool

Findings relating to the effectiveness of a current infant app (Teachbrite) supported the potential of future app-based tools for remote collection of developmental cognitive data from a diverse sample of families.

Analysis of sociodemographic measures did not reveal any significant differences between participants recruited to the study who did and did not actively use the Teachbrite app, indicating that these may not have played a role in whether parents did or did not engage more deeply in this study. A large proportion of parents reported being in either part-time or full-time work, with Hollingshead rankings indicating that occupation levels were generally high. Parents had mostly completed a relatively high level of education with most parents across samples having reached higher education. There was slightly more distribution across levels of household income, though there were still substantial proportions of parents on very high incomes. Such distributions indicate that, even when using a remote data collection tool specifically aimed at recruiting a diverse sample, developmental research may typically recruit a sample skewed towards higher SES families. This was additionally supported by findings from a cluster analysis of SES variables, which found two cluster groupings largely relating to low and high SES, but with considerably fewer participants in the low SES cluster. It is important to report descriptive sociodemographic data despite null findings between groups such that emphasis is placed on limits to the generalisability of findings (Garcini et al., 2022); given the skew of the current sample, it should not be assumed that findings in the current chapter necessarily generalise to other demographic profiles, though they are useful in provide insights about this specific group.

In analyses to assess the validity of measures collective by the Teachbrite app, app scores were found to be closely related to cognitive score and infant age, with cognitive ability and age also closely related. That older infants had higher cognitive scores was in line with expectations and validated the use of the ASQ questionnaires collapsed across ages as it was in the current study. Whilst this demonstrates that the ASQ questionnaire could be used to effectively gather developmental data remotely, it should be noted that the Teachbrite app holds potential which the ASQ questionnaire does not. Though both are scalable, the prospect of future personalisation in the Teachbrite or a similar app could not be achieved by the ASQ questionnaire. Such individualisation could enable cognitive assessments to be more efficient in identifying a child's abilities, thereby potentially reducing the impact of attentional skills or temperament on completion of assessments and might also facilitate detection of differences in cognitive profiles across children from different backgrounds. Individualisation in the Teachbrite app might additionally be helpful for engaging families in research, as dissemination of research could be made specific to a child's individual developmental profile. This is in line with findings from the focus group study, in which parents indicated that the aspects of development they were most interested in was motivated in large part by their child's interests or developmental stage. Further to opportunities for individualisation that are afforded by the Teachbrite app but not the standard ASQ questionnaire, an app-based tool also holds potential for the integration of other cognitive tasks, such as those that children complete themselves. For example, the iTapp toddler app which was demonstrated to parents during focus groups contained touchscreen tasks for toddlers to complete themselves as well as activities like in the Teachbrite app. These tasks were designed by the 'Pip and the Brain Explorers' at Kings College London (http://oneofakindcharity.com/_pages/pip/) and were designed to assess abilities relating to rate of learning and inhibitory control (a form of EF), thereby further demonstrating the potential for an app to also assess ability in specific cognition domains.

Strong positive associations between ASQ and app scores indicated that the Teachbrite app is a valid tool for collecting infant cognitive data, whilst high cohesion between two app score measures provided reassurance that rankings are suitable and reflect true difference in task difficulty. Whilst relations between app score and age did not reach significance after excluding participants who only provided valid data for a small number of tasks, this is likely due to limited power caused by small sample size and associations showed a strong positive trend. That positive associations were found between cognitive and app scores even when controlling for the number of valid trials consolidates that app scores may be informative even when the app is used in low frequency. These findings provided support for the use of the Teachbrite app in collecting valid cognitive data remotely from infants between 4- and 18-months-old.

5.4.3 Aim three: relation between SES and cognitive profiles

As mentioned, a cluster analysis involving a range of SES variables found two clusters which broadly related to a low (cluster group 2) and high (cluster group 1) SES group, which is line with findings in chapter, though there were many more members of the high versus low group. Comparisons between the two clusters found that the high SES group had higher scores on each of household income, occupation level, education level and self-rated ranking in the community. That no group differences were found for self-rated ranking in the UK and people per bedroom may suggest these measures are not as closely tied to other features of SES; in particular, it is interesting that how parents perceive their standing in the UK is not associated with more objective measures of this. Despite a wealth of work indicating an association between SES and cognition (Duncan & Magnuson, 2012; Fernald et al., 2013; Wild et al., 2013), the present study did not find any significant SES-related differences in cognitive measures. Whilst this may be due to a true lack of differences in this age range, these null results could be to a lack of statistical power caused by the small size of the low SES group. To enable strong conclusions about the relation between SES and cognitive development to be made, more diverse samples are needed so that associations across the whole range of environmental experiences can be fully explored.

5.4.4 Limitations

Though this chapter is focussed on methods to improve diversity of participation in developmental cognitive neuroscience, it remains limited by skewed representation. Focus groups attempted to recruit a range of parents through involvement with a preschool and via other local community links, though this study likely attracted parents who were already somewhat engaged in developmental work. Indeed, this is supported by focus group data, with interest in this specific study and experience of research arising as key motivations for parents' participation. As discussed, the Teachbrite app study also recruited a skewed sample of participants, and it is perhaps likely that only the most engaged parents provided plentiful app data. This limits the generalisability of findings and may cause findings to be missed due to insufficiently diverse samples (Garcini et al., 2022). In addition, whilst the current work provides information about how future work might improve diversity in research, this is only a first step and does not cover all aspects of this. Focus groups, as conducted in the current study, demonstrate possible uses of research which engages community members, though it is important to avoid tokenism (Carr et al., 2006) and to have continued engagement through which equal and lasting partnerships can be built (Garcini et al., 2022).

Limitations also exist in relation to Teachbrite app study, with issues relating to measures and methods used. Although app scores are closely related to ASQ scores, they are derived based upon task rankings which allow for an element of objectivity. This app was designed to gather data about how performance on different tasks is related and it will gain power when algorithms can be implemented which utilise upon this rather than using a more traditional approach, though the strong associations with ASQ are nonetheless encouraging. ASQ scores themselves may be limited by their nature as a parent-reported questionnaire measures, though these have been commonly used as a measure of cognition in young children across cultures (Agarwal et al., 2020).

5.4.5 Conclusion

Despite research indicating a relation between children's early experiences and cognitive development (Bradley & Corwyn, 2002; Ursache & Noble, 2016), much developmental cognitive neuroscience research is limited by a lack of diversity in participant samples. This inadequate representation limits generalisability of findings and may cause important relations between experiences and development to be missed (Garcini et al., 2022). The current chapter focussed on the development of a scalable app-based measure of early development which could be used by researchers to remotely collect cognitive data about young children. Focus groups and questionnaire data helped to identify factors which parents consider important for research utilising an app-based tool, which could have influential implications for future development of this research. Measures collected by a current app-based tool revealed strong relations between app measures, age, and other cognitive measures, supporting the validity of this tool for developmental cognitive data collection. Finally, a data-driven approach found two clusters among numerous SES variables which mapped to a low and high SES group as is typically used in research, though analyses did not reveal a relation between SES grouping and cognitive ability. These findings are together formative about methods for improving diversity of representation in future developmental cognitive research such that the relation between children’s early experiences and cognitive development may be further explored.

6 GENERAL DISCUSSION

The overarching goal of this thesis was to develop tools and expertise for increasing the diversity of toddler samples in developmental cognitive neuroscience (DCN) such that the relation between early experiences and development can be fully understood. In addition to having a methods-related goal, this thesis also focussed on increasing conceptual understanding about how socioeconomic status (SES) and development are related in infants and young toddlers. With much previous work having focussed on infants or older children, the work in this thesis provided insights about the difficult toddler age. As there is evidence of a relation between SES and development already apparent in preschoolers, this thesis explored similar in a younger age, approaching this from an adaptive framework. Given the rapid brain development which occurs during the first years of life and the considerable potential for environmental influences on this (Tierney & Nelson, 2009), functional brain development was also investigated in this thesis. To gain a greater understanding about neural measures which underpin development, the studies in this thesis have assessed the reliability of commonly used EEG measures, extended knowledge to the toddler age range, and have also explored the validity and feasibility of methods specifically focussed on including a diverse range of toddlers in DCN research.

6.1 SUMMARY OF MAIN FINDINGS

6.1.1 Chapter 2

This chapter established that a battery of eye-tracking tasks was a feasible way to assess visual attention in 18-month-olds. Two factors relating to social and non-social attention were found to explain performance across eye-tracking tasks, in line with theoretical models of visual attention. Investigations of SES measures found that a data-driven approach revealed two SES groupings mapping to typical low and high SES categories, with evidence of different patterns of visual attention between the two groups. These groupings are in line with work which indicates generally strong associations between different SES measures, though the lack of group differences for all variables supports that measures are not entirely interchangeable (Braveman et al., 2005).

Analyses using SES clusters found no difference between social and non-social attention for the high SES group, but evidence of relatively lower social and higher non-social scores for the low SES group. More specifically, the low SES group looked more to faces, were slower to find a hidden object in a working memory paradigm and were slower and less accurate in the learning phase of a cognitive control. Evidence supports that these may be adaptive differences which have arisen in response to early experiences; considering profiles of visual attention across tasks may provide insights about this which an individual task approach cannot. These findings additionally point to specific areas of attention that may benefit from future focus around the development of support for children with poorer attention skills, which may be particularly pertinent for children from low-resource backgrounds.

6.1.2 Chapter 3

Investigations into a wearable neuroimaging system which may be particularly suited for use with toddlers in field and community settings found differences across frequency bands, brain regions and video conditions which were informative about neural measures that underpin cognitive processing and added to somewhat limited knowledge about the distribution of theta and alpha power in young preschoolers. Specifically, this work found that theta power was consistently higher than alpha power and theta power was greater in posterior versus frontal regions, which adds to existing knowledge about how dominance in these frequencies shifts over development. That theta and alpha power were greater during social versus non-social conditions and there were mixed findings around theta and alpha increased over the course of video viewing contribute to understanding of the relation between these neural measures and cognitive processes, extending existing knowledge to the toddler/ preschool ages.

This chapter additionally found considerable variation in the test-retest reliability of a range of measures which differed across factors such as power type (absolute or relative), frequency band and brain region. Measures of relative theta over the whole head and over a sixty-second viewing period were the most reliable, with mostly excellent reliability. In general, absolute measures were less reliable than relative measures and measures across the whole head were generally more reliable than the difference between brain regions. Whilst it is reassuring that many measures showed moderate to excellent reliability, the considerable disparity in reliability scores signifies the importance of careful choice around which measures to use in future work.

6.1.3 Chapter 4

In further investigation of a wireless EEG system which may enable neuroimaging data collection during naturalistic interactions, chapter 4 found high acceptability and feasibility of a less-controlled design for use with toddlers. In work which demonstrates the possibility of collecting valid EEG data from young children during a free play paradigm, analyses found that relative alpha power was lower during toy exploration compared to during bubble blowing, social and non-social interactions, whilst theta power was higher during social and exploration versus bubble blowing and non-social conditions. These findings were in line with predictions and add to existing work which indicates a role for alpha in inhibitory control and theta in active learning. Significant relations between depth (length) of toy exploration and power found for each of the theta and alpha frequency bands in posterior regions only are interesting but require further investigation to be fully understood, whilst indication of a positive relation between experience of chaotic environments and depth of exploration was found but should be considered with caution as this was on the verge of significance. A positive relation may be at odds with existing work which indicated that experience of adversity was related to reduced exploration in the context of previous institutionalism but might be understood when interpreted under an adaptive framework of early experiences. Specifically, children who experience more chaotic and unpredictable environments may experience fewer opportunities and may learn to actively seek these out through greater exploration of their environment; if greater exploration is related to more distributed attention, this may also link to other findings about experience-related differences in attention (such as in chapter 2).

6.1.4 Chapter 5

This chapter investigated parental views, validity, and feasibility of a scalable app-based measure for DCN research. Findings found differences in parental views about child development research that may relate to differences in their own experiences, highlighting the need for researchers to understand specific communities that their work relates to. This work additionally discovered that individual differences and age-related developmental stage were key factors impacting parental views, whilst parents also commonly reported that an aspect of personalisation would be an important factor impacting the likelihood of them using an app-based tool. A key finding was that parents felt there was a lack of support and readily available information from the birth of their child until school or preschool age, highlighting a clear need and indicating a significant gap in information sharing which a research tool of this kind might help to fill. Parents generally indicated a willingness to share personal information in a research tool, provided the motivations for collection of this data were made clear. These findings provide insights about reasons why parents do or do not typically engage in research which might be applied to future work aiming to increase diversity of participation, whilst differences between groups of parents emphasise the need for researchers to recognise differences and understand specific perspective of communities they hope to engage.

This chapter additionally found that measures collected via an infant app were closely related to both age and cognitive measures collected via parental questionnaire, indicating that this could be a valid method for collecting cognitive data remotely about infants. That this study recruited a relatively high SES sample despite using remote methods which may be suited to broader recruitment indicates a need for researchers to make active efforts to increase diverse participation. This was also supported by uneven group sizes discovered via a cluster analysis of SES measures, which found two groups generally relating to high and low SES (similar to in chapter 2) but with far fewer participants in the low SES group. It is difficult to say whether the lack of SES-related findings in cognitive measures reflect a genuine lack of association or if this may be due to a limited sample in the current study.

6.2 CONTRIBUTIONS AND CONSIDERATIONS

6.2.1 Attentional processes

The current thesis contributes to knowledge about attentional processes in young toddlers. In chapter 2 a large sample of young toddlers (18-month-olds) were found to provide robust and good quality eye-tracking data across a battery of tasks, which supported its use as a feasible method for assessing visual attention in this age range. That this battery additionally found evidence of two latent factors largely mapping to social and non-social attention, which were further used to investigate SES-related differences in visual attention, demonstrated the potential use of this method for investigating the underlying structure of visual attention and considering how this may relate to differences in early experiences

6.2.1.1 SES and attention development

In consideration of relations between SES and attention, the current thesis has provided insights about how different SES measures relate to one another as well as measures of early visual attention and cognitive ability.

In two separate studies, data driven approaches found that SES measures typically indicated two main groups relating to traditional low and high SES backgrounds. Whilst no evidence of SES-related differences in cognitive ability were found, SES-related group differences in profiles of visual attention were discovered. Many potential reasons for this exist, not least that studies measured different domains, had different sample sizes, used different methods, and included participant samples of different ages. In fact, that differences were found in visual attention but not cognition may be informative about how relations between experiences and development emerge in early childhood. Specifically visual attention is often considered a precursor to later executive functions (Kraybill et al., 2019), which are thought to underpin other more general cognitive abilities at an older age (Diamond, 2013; Hendry et al., 2016). The current findings considered together could therefore provide some evidence that it is initial differences in visual attention which may later cascade into differences in broader cognition and may add to our understanding about mechanisms which underpin experience-development relations.

Whilst the ASQ-3 (Agarwal et al., 2020) as used in chapter 3 was designed for parental use and has been commonly used to assess cognition in young children, eye-tracking may provide more sensitive measures than parental questionnaires. This may mean potential subtle differences in development were better detected in chapter 2 and may explain the significant findings in that chapter but not in chapter 3. Further, whilst participants in both studies were of similar ages, chapter 2 included data from participants with an average age of 18-months, whilst this was the upper end of the age of participants in chapter 3. It may be that 18-months-old is around the age at which SES-development relations begin to emerge, whereas these may not yet be apparent at a younger age. Alternatively, and as reported in that chapter, the lack of SES-differences in chapter 3 may be due to lack of statistical power caused by a biased sample which included only a small number of children from low SES backgrounds.

It is nonetheless interesting to consider these findings together in the broader context of SES research. Notably, whilst both studies found similar low and high SES clusters, in both studies, not all SES measures showed differences between groups. Specifically, chapter 3 found that individuals' self-ranking within the UK did not differ between groups which might indicate that subjective measures may be less closely tied to objective measures. Interestingly, group differences for self-ranking in the community were found, which might reflect differences in how parents typically relate to others at local and national levels. Additionally, whilst education differences were not found between groups in chapter 2, these were detected in chapter 3. This may be due to the use of slightly different education measures in these studies (with years in education used in chapter 2 and education level used in chapter 3) or it might indicate that different SES measures may be more or less closely related in different contexts. For example, a child with a highly educated parent may be very likely to also live in a high-income household, but this may have no bearing on the level of resources in the home. Of course, it is likely and there is evidence that different SES measures are generally closely related, but differences such as those found in the current work indicate that how measures relate may differ across participant samples and support that measures are not simply interchangeable (Braveman et al., 2005). As Antonoplis (2023) outlines, different structural features of an environment (i.e. different measures of SES) are a sequence of related conditions, with experience of one condition impacting the likelihood of others also occurring. The current work might therefore be considered under this conceptualisation, with the SES groups providing some insight into how different SES features are related in different population samples.

The current thesis has also provided some support of an adaptive view of the relations between experience and development. Chapter 2 found that children of lower SES backgrounds looked more to faces, were slower to find a hidden object and were slower and less accurate in the learning phase of a cognitive control task, which might link into other work around adaptation and cognitive trade-offs dependent on prior experiences. Different patterns of face-looking tendencies might reflect different visual foraging techniques, poorer inhibitory control in a cognitive control task might be the result of a trade-off with enhanced task shifting (Ellis et al., 2017; Mittal et al., 2015) and poorer working memory may be the cost of enhanced procedural learning (Dang et al., 2016) which make sense in different given contexts (i.e. during poverty or in unpredictable environments). Chapter 4 also provided some support for a positive relation between experience of chaotic environments and exploration, which may be explained by a need for children who experience more chaotic and unpredictable environments to learn to actively seek out these opportunities through greater exploration of their environment as they are less likely to be provided than in less chaotic households. Of note, this latter finding was on the verge of significance therefore should be considered with caution, but nonetheless this thesis may add to a growing body of research supporting that SES-related differences are useful adaptations to experiences, rather than deficits caused by disrupted development (Ellis et al., 2017).

6.2.2 Brain functioning

In addition to measures at the cognitive level, it may be necessary to study brain development to fully understand SES-development effects. Given previous evidence of SES-related differences in neural functioning in the alpha and theta frequency bands, the current thesis focussed on measures of this during toddlerhood.

The neuroimaging work in chapters 3 and 4 made useful findings about theta and alpha power, building on, and extending existing literature about the distribution and function of neural activity in these frequencies in toddlers and preschoolers. That theta power was consistently higher than alpha power and theta power was greater in posterior versus frontal regions are quite general findings yet contribute to considerations about how dominance in these two frequency bands might change over development. Specifically, higher relative theta to alpha power in chapter 3 might indicate that alpha dominance is not yet apparent at 3- to 4-years-old but may occur later in development; this is in line with suggestions that this shift occurs at around 7-years-old (Cellier et al., 2021). Work in this thesis also found evidence of greater power, as well as significant associations between power and other measures (i.e. exploration), which were specific to posterior regions. These findings may be considered as adding to existing work which implicates posterior alpha in the maintenance of centrally-focussed attention (Orekhova et al., 2001) and supporting a shift from primarily frontal to more widespread (or even predominantly posterior) theta from infancy to early childhood (Cuevas et al., 2012; Orekhova et al., 2006).

Condition comparisons revealed that theta power was higher during social versus non-social video viewing, and during toy exploration and live social interactions compared to during bubble blowing and live non-social interactions. These findings are in line with previous findings (Jones et al., 2015; Orekhova et al., 2006) and strengthen a body of work indicating that theta may play a role in information processing (Guderian et al., 2009; Meyer et al., 2019). In particular, higher theta power found during free exploration in chapter 4 might be interpreted as further supporting the hypothesis that theta oscillations may play a role in establishing optimal conditions for information processing (Begus & Bonawitz, 2020) and provides, to my knowledge, the first such evidence in toddlers. In chapter 3, alpha measures followed the same pattern of differences as theta measures, with higher power during social versus non-social videos as predicted, however chapter 4 found that alpha power was lower during toy exploration compared to all other conditions, with no evidence of any other condition differences. These latter findings were in line with work which indicates a role for alpha in inhibitory control, with lower alpha power during toy exploration possibly reflecting more widely distributed attention (i.e. requiring less inhibition) in the current work, though the studies in chapters 3 and 4 indicate different findings about alpha during social and non-social conditions. It is possible these findings are due to differences in paradigms, with one study using pre-recorded videos and the other using live researcher interactions. In these results, videos induced higher power to a social compared to non-social condition which live interactions did not. Whilst this may appear counterintuitive at first, it might be explained by challenges associated with controlling conditions in live paradigms. Specifically, in non-social pre-recorded videos, the non-social condition showed mechanical toys moving; these toys require initial human input to establish movement, however this was cropped out of videos, meaning the condition displayed no human aspect. In live interactions, children saw a researcher move the toys to the table and make the initial movement required to start the toy. Though researchers were instructed to minimise eye-contact (in comparison to the social condition in which they sung nursery rhymes and actively tried to engage the child), it may be that toddlers did not interpret this condition as entirely non-social and that the presence and contribution of a researcher induced neural functioning more akin to during social processing. Further work might use a range of conditions which involve varying degrees of social input to investigate whether and how alpha power varies in relation to social processing. Given interpretation of lower alpha power as reflective of higher inhibitory control, differences in alpha power across paradigms and conditions might not be due to social input per se, but rather the degree of inhibition that is involved, therefore paradigms which vary in the amount of inhibitory control required may also be informative. Further to paradigm differences, the different results in chapters 3 and 4 may be indicative of developmental change over the toddler years with an average age of 33 months in chapter 4 compared to 38 months in chapter 3. Although this age gap is not large, it may cover a significant period of development during which there are considerable changes in neural processing; collapsing across these ages and looking at age effects might aid understanding of development over this period. Further findings relating to changes in power over time and considering associations between power and toy exploration provide further insights about neural processing in the toddler and preschool ages but require further exploration to be fully understood.

Whilst there is some existing literature about theta and alpha during toddlerhood, the current thesis adds to this somewhat limited body of work, providing insights about neurocognitive development and processes which may underpin skills in other domains. Given the theorised links between these measures (theta and alpha), exploration and attention, and work indicating that visual attention may be an early domain in which SES-development relations can be observed (as in chapter 2), measures of theta and alpha may be sensible candidates for future work aiming to explore the early impact of experiences on neural development.

6.2.3 Contribution to gap in existing literature

In addition to the findings already discussed, the current thesis has additionally provided tools and expertise for methods to improve diversity and include a greater number of toddlers in DCN research. Methodology development has included assessment of acceptability, feasibility, reliability, and validity of several methods and experimental designs which may be particularly suited for use with toddlers, and which may make research more accessible for a greater number of families.

6.2.3.1 Scalable cognitive methods

One way to improve inclusion in research is to use methods which can reach a large number of families. This thesis utilised two types of methods which may be well-suited for such large-scales studies, both of which returned positive results indicating their potential for this work. In chapter 2 a large sample of young toddlers (18-month-olds) were found to provide robust and good quality eye-tracking data across a battery of tasks, which supported its use as a feasible method for assessing visual attention in this age range. Whereas chapter 2 had already utilised an eye-tracking battery in a large-scale study, chapter 5 investigated the validity of an app-based method in a much smaller sample. This work provided insights about how parents might use an app-based tool in future work and additionally validated an existing infant app (‘Teachbrite’) for collecting cognitive data remotely from a sample of parents of infants. Cognitive scores derived from parent-reported performance on researcher-developed tasks correlated highly with age and cognitive scores from online questionnaires, thereby indicating the suitability of the ‘Teachbrite’ app for collecting neurocognitive data in this way. Notably, app-based tools can be used entirely remotely, therefore reducing many barriers associated with families physically attending a setting to participate in research. They are additionally low-cost and well suited to being scaled-up, meaning this method might be suitable for reaching many families. Both these factors mean app-based methods may lend themselves to including a more diverse range of families in DCN research, with findings in the current thesis supporting such use.

Nonetheless, there are weaknesses associated with both app-based and eye-tracking methods which ought to be considered by future researchers designing their research. Though eye-tracking is often cited as a more direct measure of individuals’ attention, this relies on assumptions about underlying processes (which are discussed in greater depth in section 6.3 below) and may be impacted by environmental factors such as lighting (Aslin, 2007). This latter point may be particularly pertinent when considering how eye-tracking methods may be scaled up for use in a variety of settings and care should be taken to control or adjust for environmental differences in experimental setups. Because the current work used a screen-based eye-tracker, there may be challenges around expanding this battery for use with older toddlers (i.e. 2-year-olds) due to rapidly changing attentional and physical abilities (see intro 1.3 for discussion); it may be that wearable or other eye-tracking systems may be better suited for scaling such research across a broader age range. In addition, use of eye-tracking typically requires the presence of a researcher or technician, who can set up the system for each individual participant, run the task battery and troubleshoot any technological difficulties, meaning that scalability is at least somewhat limited. Whilst it may be possible to train others to perform this data collection role, there would still be a limited capacity based upon the number of trained people and eye-tracking systems available, meaning scalability is lower than for app-based methods, which may be freely used by participants themselves.

Whilst apps are advantageous over other methods in this respect, meaning they may also be used to collect data more frequently, this also could lead to challenges with app-collected data. As data collection is completely remote and requires no interaction directly with a researcher, it may be difficult to ensure consistency across data collection and to ensure that app-users do indeed fit criteria for a particular study. Whilst it might be possible to build in identity or attentional checks, app-based methods ultimately rely on trust that participants largely use the tool correctly and accurately, and that any biases in responses may be systematic such that it impacts data from all participants similarly. A combination of parent-report and touchscreen tasks which toddlers complete themselves may be advantageous as both data types could be contributed to understanding a child's cognitive profile, though again it is difficult to control for outside factors (i.e. presence of other siblings, TV turned on, etc.) which might impact responses. This may be a challenge which ought to be considered in much research which uses scalable measures, since it is perhaps likely that methods need to be used in multiple settings to effectively reach a wide range of families.

6.2.3.2 Mobile EEG

In addition to the development of cognitive methods, this thesis included advancements in the use of a portable EEG system for measuring development at the neural level. This work focussed around assessing the feasibility and reliability of a low-density and low-cost set-up which could be used to collect high quality neural data from toddlers in a range of field and community settings.

Test-retest reliability of a range of EEG measures collected from 2.5- to 4-year-olds during a screen-based paradigm ranged from poor to excellent, indicating considerable variability. Relative power measures were generally more reliable than absolute measures, and measures averaged across brain regions were more reliable than differences between regions, providing some general guidelines about the reliability of such metrics. More specifically, relative theta power over the whole head and over a sixty-second period of video viewing offered the highest reliability, with mostly excellent reliability scores. Analogous measures were also the most reliable metrics of alpha power though these were lower than for theta, with largely moderate reliability scores. It is important to point out that it is possible that reliability could differ depending on the age of participants, the experimental design, or the EEG system in use, meaning the current findings might not necessarily relate to other set-ups. Nonetheless, these findings are informative about the test-retest reliability of several measures which might reflect broader patterns of reliability and highlight the importance of considering reliability when choosing measures for future research. A key point which arose from these findings was the large variability in test-retest reliability scores of different measures, which emphasises the potential impact choice of measure could have on the reliability of findings and stresses need for careful consideration of this.

Acceptability and feasibility of a less-controlled design for collecting neurocognitive data from toddlers was found to be high in chapter 4, indicating the potential for portable EEG systems to increase research in this age range. That this work collected usable EEG data during live interactions and a free-play session additionally signifies possibilities for collecting neural data during more naturalistic experiences, which may improve ecological validity of studies and enable researchers to further explore how children interact with the world. The free-play design and use of a low-density, portable EEG system in the current work might provide evidence of a feasible study design that could be easily translated to different environments such as field and community settings, which might therefore facilitate inclusion of a wider range of families in this research than lab studies typically enable.

One method which might be used specifically to increase the diversity of participation in DCN research is an element of community engagement. The current thesis explored the first stage of this by conducting focus groups, designed to understand parental views about child development research and how they might use an app-based tool with their children. As mentioned, this work indicated that parents feel there is a lack of support and information from birth until their child starts at school or preschool, indicating a particular gap which researchers might be able to fill. Additionally, it was found that an app-based tool should be personalised to suit individual children and their particular stage of development, whilst other findings show considerable variability among parental views, signifying the importance of engaging members of specific communities in research to fully understand different perspectives. Employing such principles could help increase diversity of participants in DNC work, thereby improving representation and generalisability of findings.

The work in this thesis has provided information and guidance about how a range of methods may be utilised in future work to increase the participation of toddlers from diverse backgrounds in DCN research. This includes measures at both the cognitive and neural levels and methods which can be used entirely remotely or in numerous environments including community or field settings in addition to utilising community engagement as a means of engaging community members in research.

6.3 RETURN OF RESEARCH RESULTS TO PARTICIPANTS

Whilst this thesis discussed methods which may include informing parents about their children's development, it is important to consider ethical implications of such. This so-called 'return of results' to participants remains under debate, with some arguing that it would be unethical to provide participants with information about themselves which could be clinically-relevant or helpful, and others arguing that it is more ethical not to share information which could be potentially misunderstood or misused (Burke et al., 2014).

There is an interplay between research and clinical care whereby research may be clinically-relevant, yet the two are also distinct in important ways. Whereas research is focussed on producing generalizable knowledge, clinical practice is focussed on optimising health care for individuals (Burke et al., 2014). Consequently, a clinician's primary role is to work in the best interests of the patient, whilst a researcher's focus is on maintaining research integrity. Further, whilst individuals have a legal right to receive information about themselves in a clinical setting, the same laws do not apply to research (Burke et al., 2014). Even when research is not related to clinical care, there are arguments for and against dissemination of information to participants. On one hand, sharing of information may show respect and gratitude to participants who have volunteered their time to take part in research, and may lead to greater engagement in research (Downey et al., 2018). On the other hand, demonstrating respect for participants may be achieved by other means, such as providing clear information about return of results, and sharing generalised findings from studies participants have taken part in (Downey et al., 2018). Findings from non-clinical research may be informative for health or lifestyle management, or for psychological well-being (Ravitsky & Wilfond, 2006), however returning results could equally cause undue distress and may actually have a negative impact on well-being (Ashida et al., 2010; Dixon-Woods et al., 2011). Perhaps important is that research is done to provide knowledge and, as such, results are often provisional in nature and are not always robust measures at the individual level.

In the context of child development, there is an additional layer of complexity to consider with regards to return of results. It is necessary to balance the rights and best interest of the parent and the child, which may sometimes conflict with one another (Holm, 2017). Under the 'best interests of the child' framework, decisions about children's participation and information in research are shared between parents, researchers, the child themselves and, on occasion, the State (Zawati et al., 2014). Returning results about how environment or family background relates to children's development may be particularly sensitive. Even when disseminated in general, not individual, terms, care should be taken that findings do not support a rhetoric whereby parents who experience socioeconomic disadvantage are blamed for differences in their child's abilities, but should instead support and empower parents to support their child's development (Blakey et al., 2024). Given the factors discussed here, it is clear that whether results are disseminated to participants should be carefully considered and, where this is done, that considerations are weighed-up and dissemination is done in the most ethical way.

6.4 LIMITATIONS AND FUTURE DIRECTIONS

While the current set of studies did have some limitations, these provide interesting opportunities for future research.

Although this thesis has provided insights about how SES measures relate to one another and to other measures of development, findings are limited by the specific SES measures used. There is an abundance of factors of early environments which might be important to investigate, but only a small number are included in the studies here. Even when specifically considering factors that are traditionally thought of as SES (without including additional features such as noise, pollution, etc.) there are numerous elements which could be included. Chapter 1 here included measures of education, occupation, resources in the home and postcode-based deprivation, whilst chapter 5 considered household income, occupation, education, people per bedroom and self-rated ranking in both the community and the UK. Whilst these different features may be closely related, there are hypotheses and evidence to support that they reflect distinct properties of individuals' environments and are not interchangeable (Geyer et al., 2006). Choice of SES variables might therefore impact findings, meaning it is difficult to make broad statements from the current work about more general environment-development relations. As suggested by Antonoplis (2023), there may be use in studies looking at the effect of individual SES features (i.e. income or occupation) on developmental outcome and interpreting these as such (rather than in terms of SES), such that a body of work is built around the relation between separate features and development (rather than SES and development), as exists for associations between sex/gender and development. Some researchers use an approach to quantifying early environments that is different to human-defined SES constructs and instead characterises environmental properties through data-driven methods. Such work might use wearable and portable technologies to measure properties of an environment (e.g. noise levels or predictability of parent behaviour in a home), which are then considered in relation to a child's development (i.e. in Wass et al., 2019). Whilst this approach is different to traditional consideration of SES, it may be informative about mechanisms underpinning observed associations. In addition to potential challenges around how SES was defined and measured in this thesis, interpretation of SES findings within an adaptive framework as in the present work is a theoretical perspective which is difficult to empirically prove. Future work might benefit from considering differences across tasks at the individual level to further explore this.

Measuring neural activity can offer great insights about development at a different level of processing to traditional behavioural measures and may point towards potential mechanisms which underpin behaviours however is limited by use of assumptions about what measures reflect. That is, higher EEG power is generally interpreted as greater activity and engagement of a particular brain region, but questions remain about what information oscillations actually carry. Understanding these mechanisms fully requires causal knowledge about how neural functioning relates to behaviours, but this is challenging in cognitive neuroscience (Danks & Davis, 2023). A lack of causal knowledge limits the development of optimal support and interventions, as these must act on causes rather than effects of the phenomenon it is hoped to alter (Danks & Davis, 2023), therefore it is crucial that future work focusses on building causal knowledge if research is to ultimately help all individuals achieve their optimal development (World Health Organization, 2022).

Mechanistic limitations are also somewhat true of work utilising eye-tracking, whereby interpretations of looking behaviour often rely on assumptions about the underlying processes. Two key assumptions are that there is a relation between eye movements and cognition at all (i.e. that eyes on a stimuli equals processing) and that processing of a stimuli begins immediately upon fixation of it (Cullipher et al., 2018). Other assumptions are made about specific eye-tracking measures, such as that longer looking times to one stimulus over another reflects ability to discriminate between the two, but it is possible that measures and individuals' looking patterns are impacted by other factors including lighting levels, physical features of stimuli and personal interests or that patterns may be explained by other cognitive processes (Aslin, 2007). Challenges may be overcome by careful experimental design and analysis choices but should nonetheless remain a consideration of researchers interpreting such work.

With both eye-tracking and neuroimaging methods, there is considerable processing which must be done before data can be readily interpreted and choices around this, as well as which measures to use, can also have an influence on findings. Moves are being made towards more consistent processing of data, for example with the development of various processing pipelines for EEG data (Gabard-Durnam et al., 2018; Haartsen et al., 2021), and new approaches are being developed which may enable researchers to optimise individual differences and deal with noisy data in a novel way (Gui et al., 2022). Such methods could additionally help with often considerable variability in the amount of data collected from different children (Gui et al., 2022), which can cause difficulties in interpreting whether findings may reflect differences in temperament or other factors which are somewhat masked by this variability. Some efforts were made to control for differences in data quantity in this thesis but remain a limitation of cognitive and neuroimaging work.

Whilst the current thesis provided information and guidance for including a greater number of toddlers in DCN research, a small number of children did not tolerate demands of the studies. Though particularly true for neuroimaging work, it is possible that this could also impact cognitive methods, perhaps due to factors including confidence, sensory burdens, or interest in specific activities. Methods developed here – such as those that can be used in the home or settings familiar to the child and those that enable the child some choice of exploration - may help overcome such issues, however it may be likely that some children will always refuse to wear a neuroimaging cap. Though a small number of refusals may not be of great concern, it is necessary that systematic exclusions are recognised and adaptive efforts made, and it is particularly important to note this limitation when considering inclusivity in research.

Though a key topic of this thesis was methods for increasing diversity and improving generalisability of DCN research, some studies' participant samples were biased towards high socioeconomic families. On the one hand this demonstrates the need for work which focusses on this and provides an opportunity to demonstrate how results should only be interpreted with regards to the population it represents; on the other, this limits generalisability of the present work and suggests that much more effort is needed to achieve more diverse participation in DCN work. Insights provided by focus groups indicated one possible route for building a tool to engage more families, though it is likely that continued engagement as well as involving other specific communities that researchers would like to engage are needed for success. Somewhat similarly, the less-controlled paradigm in chapter 4 utilised portable EEG systems and demonstrated methods which could be used in field and community settings but was limited in its considerations of how to move this from the laboratory to the field. Other research has focussed on this, with evidence indicating that use of tablet, EEG and eye-tracking methods may be feasible and well-accepted in home and low-resource settings (Bhavnani et al., 2019, 2022; Lockwood Estrin et al., 2022, 2023; Troller-Renfree et al., 2021). One general limitation of this thesis is that it mostly provides first steps, recommendations, and demonstrations about how methods might be used to increase diversity and toddler participant in DCN research without necessarily conducting the large-scale studies which utilise this work; nonetheless, perhaps the greatest contribution of this thesis is the basis it provides for future research to do so.

6.5 CONCLUDING REMARKS

There is huge variability in the experiences which individuals might have at both local (i.e. even within a small community) and global levels which can have a significant and lasting impact across a range of domains throughout life. To fully understand the complexities of relations between experiences and development, research must include individuals from diverse backgrounds, as without representative samples that include individuals from all backgrounds this can never be fully understood. Whilst many researchers acknowledge the problem of lack of diversity in research, few have focussed on this and made real efforts to improve. This thesis made diversity a key aim to help fill this hole and has made meaningful contributions about how researchers might utilise methods and use different approaches to increase diverse participation. As well as providing expertise for improving diversity, this thesis has provided useful contributions to existing understanding about the relation between SES and neurocognitive development in young toddlers. It has found evidence that data-driven SES groupings may be associated with differences in profiles of visual attention (which may be precursors to later SES differences in executive functioning) in toddlers, but perhaps not with more general cognitive ability in young infants. Specifically, children who were deemed to be from lower SES families looked more to static images of faces, possibly indicating differences in visual processing styles, and were slower and less accurate in a task that relies on inhibitory control. These findings might suggest that processing of social stimuli with limited input and inhibitory control are cognitive domains that are particularly sensitive to SES effects and may play a mechanistic role in broader SES-related differences later in life. More research is required to further explore these associations and how this information might be used to help all children reach their optimal development, however the current thesis contributes to existing knowledge about how SES and development of attentional processes are related during early toddlerhood.

In addition to conceptual advances, this thesis offers guidance for study designs which can gather good quality and high quantity neurocognitive data from toddlers and provide insights about neural measures underpinning cognition in toddlerhood. This thesis employed various methods and utilised technical advances to develop and assess study acceptability, feasibility, reliability and validity of designs and approaches for improving the inclusion of toddlers from diverse backgrounds in DCN research. It has made both theoretical and practical contributions of expertise and knowledge which will facilitate improved future research investigating the impact of experiences on early development.

There is a need to understand relations between early experiences and development such that the impact of inequalities is diminished, and so that all individuals can be supported to reach their optimum. Whilst the contributions of a single thesis cannot hope to solve such a huge and complex matter, the work in the current thesis provides blueprints and guidance from which future research into this can be built, and thus plays a small role in the equalling of all children’s development.

References

Abels, M., & Hutman, T. (2015). Infants' behavioral styles in joint attention situations and parents' socio-economic status. *Infant Behavior and Development*, *40*, 139–150. https://doi.org/10.1016/j.infbeh.2015.05.004

Adam, N., Blaye, A., Gulbinaite, R., Delorme, A., & Farrer, C. (2020). The role of midfrontal theta oscillations across the development of cognitive control in preschoolers and school-age children. *Developmental Science*, *23*(5), e12936. https://doi.org/10.1111/desc.12936

Agarwal, P. K., Xie, H., Sathyapalan Rema, A. S., Rajadurai, V. S., Lim, S. B., Meaney, M., & Daniel, L. M. (2020). Evaluation of the Ages and Stages Questionnaire (ASQ 3) as a developmental screener at 9, 18, and 24 months. *Early Human Development*, *147*, 105081. https://doi.org/10.1016/j.earlhumdev.2020.105081

Ahmed, A., & Ruffman, T. (1998). Why do infants make A not B errors in a search task, yet show memory for the location of hidden objects in a nonsearch task? *Developmental Psychology*, *34*(3), 441–453. https://doi.org/10.1037/0012-1649.34.3.441

Alcauter, S., Lin, W., Smith, J. K., Short, S. J., Goldman, B. D., Reznick, J. S., Gilmore, J. H., & Gao, W. (2014). Development of Thalamocortical Connectivity during Infancy and Its Cognitive Correlations. *Journal of Neuroscience*, *34*(27), 9067–9075. https://doi.org/10.1523/JNEUROSCI.0796-14.2014

Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R., Langen, J. van, & Kievit, R. A. (2021). *Raincloud plots: A multi-platform tool for robust data visualization* (No. 4:63). Wellcome Open Research. https://doi.org/10.12688/wellcomeopenres.15191.2

Amso, D., & Scerif, G. (2015). The attentive brain: Insights from developmental cognitive neuroscience. *Nature Reviews Neuroscience*, *16*(10), Article 10. https://doi.org/10.1038/nrn4025

Anderson, P. J., & Reidy, N. (2012). Assessing Executive Function in Preschoolers. *Neuropsychology Review*, *22*(4), 345–360. https://doi.org/10.1007/s11065-012-9220-3

Angulo-Ruiz, B. Y., Muñoz, V., Rodríguez-Martínez, E. I., & Gómez, C. M. (2021). Absolute and relative variability changes of the resting state brain rhythms from childhood and adolescence to young adulthood. *Neuroscience Letters*, *749*, 135747. https://doi.org/10.1016/j.neulet.2021.135747

Antonoplis, S. (2022). *Studying Socioeconomic Status: Conceptual Problems and an Alternative Path Forward*. PsyArXiv. https://doi.org/10.31234/osf.io/29v6b

Antonoplis, S. (2023). Studying Socioeconomic Status: Conceptual Problems and an Alternative Path Forward. *Perspectives on Psychological Science*, *18*(2), 275–292. https://doi.org/10.1177/17456916221093615

Arnett, J. J. (20081013). The neglected 95%: Why American psychology needs to become less American. *American Psychologist*, *63*(7), 602. https://doi.org/10.1037/0003-066X.63.7.602

Arora, S., Lawrence, M. A., & Klein, R. M. (2020). The Attention Network Test Database: ADHD and Cross-Cultural Applications. *Frontiers in Psychology*, *11*, 388. https://doi.org/10.3389/fpsyg.2020.00388

Ashida, S., Koehly, L. M., Roberts, J. S., Chen, C. A., Hiraki, S., & Green, R. C. (2010). The role of disease perceptions and results sharing in psychological adaptation after genetic susceptibility testing: The REVEAL Study. *European Journal of Human Genetics*, *18*(12), 1296–1301. https://doi.org/10.1038/ejhg.2010.119

Aslin, R. N. (2007). What's in a look? *Developmental Science*, *10*(1), 48–53. https://doi.org/10.1111/j.1467-7687.2007.00563.x

Baillargeon, R., & Graber, M. (1988). Evidence of location memory in 8-month-old infants in a nonsearch AB task. *Developmental Psychology*, *24*(4), 502–511. https://doi.org/10.1037/0012-1649.24.4.502

Baillargeon, R., Spelke, E. S., & Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition*, *20*(3), 191–208. https://doi.org/10.1016/0010-0277(85)90008-3

Begnoche, J. P., Brooker, R. J., & Vess, M. (2016). EEG Asymmetry and ERN: Behavioral Outcomes in Preschoolers. *PLOS ONE*, *11*(5), e0155713. https://doi.org/10.1371/journal.pone.0155713

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*, *113*(44), 12397–12402. https://doi.org/10.1073/pnas.1603261113

Begus, K., Southgate, V., & Gliga, T. (2015). Neural mechanisms of infant learning: Differences in frontal theta activity during object exploration modulate subsequent object recognition. *Biology Letters*, *11*(5), 20150041. https://doi.org/10.1098/rsbl.2015.0041

Bell, M. A. (1998). The ontogeny of the EEG during infancy and childhood: Implications for cognitive development. In B. Garreau (Ed.), *Neuroimaging in child neuropsychiatric disorders* (pp. 97–111). Springer. https://doi.org/10.1007/978-3-642-95848-9_9

Bell, M. A., & Cuevas, K. (2012). Using EEG to Study Cognitive Development: Issues and Practices. *Journal of Cognition and Development*, *13*(3), 281–294. https://doi.org/10.1080/15248372.2012.691143

Belsky, J., Spritz, B., & Crnic, K. (1996). Infant Attachment Security and Affective-Cognitive Information Processing at Age 3. *Psychological Science*, *7*(2), 111–114. https://doi.org/10.1111/j.1467-9280.1996.tb00339.x

Berger, H. (1929). *Über das Elektrenkephalogramm des Menschen* | *SpringerLink*. https://link.springer.com/article/10.1007/BF01797193

Bernier, A., Carlson, S. M., Deschênes, M., & Matte-Gagné, C. (2012). Social factors in the development of early executive functioning: A closer look at the caregiving environment: Caregiving and child executive functioning. *Developmental Science*, *15*(1), 12–24. https://doi.org/10.1111/j.1467-7687.2011.01093.x

Best, J. R., & Miller, P. H. (2010). A Developmental Perspective on Executive Function. *Child Development*, *81*(6), 1641–1660. https://doi.org/10.1111/j.1467-8624.2010.01499.x

Bhavnani, S., Estrin, G. L., Haartsen, R., Jensen, S. K. G., Gliga, T., Patel, V., & Johnson, M. H. (2021). EEG signatures of cognitive and social development of preschool children–a systematic review. *PLOS ONE*, *16*(2), e0247223. https://doi.org/10.1371/journal.pone.0247223

Bhavnani, S., Mukherjee, D., Dasgupta, J., Verma, D., Parameshwaran, D., Divan, G., Sharma, K. K., Thiagarajan, T., & Patel, V. (2019). Development, feasibility and acceptability of a gamified cognitive DEvelopmental assessment on an E-Platform (DEEP) in rural Indian pre-schoolers – a pilot study. *Global Health Action*, *12*(1), 1548005. https://doi.org/10.1080/16549716.2018.1548005

Bhavnani, S., Parameshwaran, D., Sharma, K. K., Mukherjee, D., Divan, G., Patel, V., & Thiagarajan, T. C. (2022). The Acceptability, Feasibility, and Utility of Portable Electroencephalography to Study Resting-State Neurophysiology in Rural Communities. *Frontiers in Human Neuroscience*, *16*. https://www.frontiersin.org/articles/10.3389/fnhum.2022.802764

Blair, C., Granger, D. A., Willoughby, M., Mills-Koonce, R., Cox, M., Greenberg, M. T., Kivlighan, K. T., Fortunato, C. K., & the FLP Investigators. (2011). Salivary Cortisol Mediates Effects of Poverty and Parenting on Executive Functions in Early Childhood: Cortisol and Cognition. *Child Development*, *82*(6), 1970–1984. https://doi.org/10.1111/j.1467-8624.2011.01643.x

Blair, C., & Raver, C. C. (2012). Child Development in the Context of Adversity. *The American Psychologist*, *67*(4), 309–318. https://doi.org/10.1037/a0027493

Blair, M. R., Watson, M. R., & Meier, K. M. (2009). Errors, efficiency, and the interplay between attention and category learning. *Cognition*, *112*(2), 330–336. https://doi.org/10.1016/j.cognition.2009.04.008

Blakeslee, S. B., Vieler, K., Horak, I., Stritter, W., & Seifert, G. (2023). *Planting Seeds for the Future: Scoping Review of Child Health Promotion Apps for Parents—PMC*. Retrieved 27 September 2024, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10401193/

Blakey, E., Outhwaite, L. A., Braithwaite, E., & Hendry, A. (2024). *How we Measure and Operationalise Socioeconomic Disadvantage in Developmental Research has Implications for Theory and Policy*. OSF. https://doi.org/10.31234/osf.io/rx9c4

Blanco, N. J., & Sloutsky, V. M. (2020). Attentional mechanisms drive systematic exploration in young children. *Cognition*, *202*, 104327. https://doi.org/10.1016/j.cognition.2020.104327

Blasi, A., Lloyd-Fox, S., Johnson, M. H., & Elwell, C. (2014). Test–retest reliability of functional near infrared spectroscopy in infants. *Neurophotonics*, *1*(2), 025005. https://doi.org/10.1117/1.NPh.1.2.025005

Bonthrone, A. F., Chew, A., Kelly, C. J., Almedom, L., Simpson, J., Victor, S., Edwards, A. D., Rutherford, M. A., Nosarti, C., & Counsell, S. J. (2021). Cognitive function in toddlers with congenital heart disease: The impact of a stimulating home environment. *Infancy*, *26*(1), 184–199. https://doi.org/10.1111/infa.12376

Bornstein, M. H., Hahn, C.-S., & Suwalsky, J. T. D. (2013). Physically developed and exploratory young infants contribute to their own long-term academic achievement. *Psychological Science*, *24*(10), 1906–1917. https://doi.org/10.1177/0956797613479974

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic Status and Child Development. *Annual Review of Psychology*, *53*(1), 371–399. https://doi.org/10.1146/annurev.psych.53.100901.135233

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Braithwaite, E., Jones, E. J. H., Johnson, M. H., & Holmboe, K. (2019). *Dynamic modulation of frontal theta power predicts cognitive ability in infancy* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/3wgxc

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? In *Progress in Brain Research*. Elsevier. https://doi.org/10.1016/bs.pbr.2020.05.007

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braveman, P. A., Cubbin, C., Egerter, S., Chideya, S., Marchi, K. S., Metzler, M., & Posner, S. (2005). Socioeconomic Status in Health Research: One Size Does Not Fit All. *JAMA*, *294*(22), 2879. https://doi.org/10.1001/jama.294.22.2879

Brito, N. H., Fifer, W. P., Myers, M. M., Elliott, A. J., & Noble, K. G. (2016). Associations among family socioeconomic status, EEG power at birth, and cognitive skills during infancy. *Developmental Cognitive Neuroscience*, *19*, 144–151. https://doi.org/10.1016/j.dcn.2016.03.004

Brooker, R. J. (2018). Maternal behavior and socioeconomic status predict longitudinal changes in error-related negativity in preschoolers. *Child Development*, *89*(3), 725–733. https://doi.org/10.1111/cdev.13066

Burke, W., Evans, B. J., & Jarvik, G. P. (2014). Return of results: Ethical and legal distinctions between research and clinical care. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics*, *166*(1), 105–111. https://doi.org/10.1002/ajmg.c.31393

Byun, S., Han, J. W., Kim, T. H., & Kim, K. W. (2016). Test-Retest Reliability and Concurrent Validity of a Single Tri-Axial Accelerometer-Based Gait Analysis in Older Adults with Normal Cognition. *PLoS ONE*, *11*(7), e0158956. https://doi.org/10.1371/journal.pone.0158956

Caldwell, B. M., & Bradley, R. H. (1979). Home observation for measurement of the environment. *Little Rock, AR: University of Arkansas at Little Rock.*

Calkins, S. D. (2007). The emergence of self-regulation: Biological and behavioral control mechanisms supporting toddler competencies. In *Socioemotional development in the toddler years: Transitions and transformations* (pp. 261–284). Guilford Press.

Canen, M. J., & Brooker, R. J. (2017). ERN, theta power, and risk for anxiety problems in preschoolers. *Biological Psychology*, *123*, 103–110. https://doi.org/10.1016/j.biopsycho.2016.12.003

Carr, S. M., Clarke, C. L., Molyneux, J., & Jones, D. (2006). Facilitating participation: A health action zone experience. *Primary Health Care Research and Development*, *7*(2), 147–156. https://doi.org/10.1191/1463423606pc284oa

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. *Journal of Personality and Social Psychology*, *67*(2), 319–333. https://doi.org/10.1037/0022-3514.67.2.319

Cavanagh, J. F., Figueroa, C. M., Cohen, M. X., & Frank, M. J. (2012). Frontal Theta Reflects Uncertainty and Unexpectedness during Exploration and Exploitation. *Cerebral Cortex*, *22*(11), 2575–2586. https://doi.org/10.1093/cercor/bhr332

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

Cellier, D., Riddle, J., Petersen, I., & Hwang, K. (2021). The development of theta and alpha neural oscillations from ages 3 to 24 years. *Developmental Cognitive Neuroscience*, *50*, 100969. https://doi.org/10.1016/j.dcn.2021.100969

Centre for Early Childhood. (2023). *Shaping Us*. Early Years - Shaping Us. https://shapingus.centreforearlychildhood.org/

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

Chevallier, C., Parish-Morris, J., McVey, A., Rump, K. M., Sasson, N. J., Herrington, J. D., & Schultz, R. T. (2015). Measuring social attention and motivation in autism spectrum disorder using eye-tracking: Stimulus type matters. *Autism Research*, *8*(5), 620–628. https://doi.org/10.1002/aur.1479

Chica, A. B., Bartolomeo, P., & Lupiáñez, J. (2013). Two cognitive and neural systems for endogenous and exogenous spatial attention. *Behavioural Brain Research*, *237*, 107–123. https://doi.org/10.1016/j.bbr.2012.09.027

Chrastil, E. R., Rice, C., Goncalves, M., Moore, K. N., Wynn, S. C., Stern, C. E., & Nyhus, E. (2022). Theta oscillations support active exploration in human spatial navigation. *NeuroImage*, *262*, 119581. https://doi.org/10.1016/j.neuroimage.2022.119581

Clarke, A. R., Barry, R. J., McCarthy, R., & Selikowitz, M. (2001). Age and sex effects in the EEG: Development of the normal child. *Clinical Neurophysiology*, *112*(5), 806–814. https://doi.org/10.1016/S1388-2457(01)00488-6

Clearfield, M., & McKenzie, K. (2013). The Effects of Socio-Economic Status on Infant Attention. *Infant and Child Development*, *22*. https://doi.org/10.1002/icd.1770

Clearfield, M. W., & Jedd, K. E. (2013). The Effects of Socio-Economic Status on Infant Attention. *Infant and Child Development*, *22*(1), 53–67. https://doi.org/10.1002/icd.1770

Clearfield, M. W., & Niman, L. C. (2012). SES affects infant cognitive flexibility. *Infant Behavior and Development*, *35*(1), 29–35. https://doi.org/10.1016/j.infbeh.2011.09.007

Colombo, J. (2001). The Development of Visual Attention in Infancy. *Annual Review of Psychology*, *52*(1), 337–367. https://doi.org/10.1146/annurev.psych.52.1.337

Colson, E. R., & Dworkin, P. H. (1997). Toddler development. *Pediatrics in Review*, *18*(8), 255–259. https://doi.org/10.1542/pir.18-8-255

Conejero, Á., Guerra, S., Abundis-Gutiérrez, A., & Rueda, M. R. (2018). Frontal theta activation associated with error detection in toddlers: Influence of familial socioeconomic status. *Developmental Science*, *21*(1), e12494. https://doi.org/10.1111/desc.12494

Conejero, A., & Rueda, M. (2020). *Developing attention and selfregulation in infancy and childhood.* (pp. 505–517). https://doi.org/10.1016/C2017-0-00830-3

Conejero, Á., & Rueda, M. R. (2018). Infant temperament and family socio-economic status in relation to the emergence of attention regulation. *Scientific Reports*, *8*(1), 11232. https://doi.org/10.1038/s41598-018-28831-x

Connor, C. E., Egeth, H. E., & Yantis, S. (2004). Visual Attention: Bottom-Up Versus Top-Down. *Current Biology*, *14*(19), R850–R852. https://doi.org/10.1016/j.cub.2004.09.041

Conroy, K., Sandel, M., & Zuckerman, B. (2010). Poverty Grown Up: How Childhood Socioeconomic Status Impacts Adult Health. *Journal of Developmental & Behavioral Pediatrics*, *31*(2), 154–160. https://doi.org/10.1097/DBP.0b013e3181c21a1b

Cooper, S. R., Gonthier, C., Barch, D. M., & Braver, T. S. (2017). The Role of Psychometrics in Individual Differences Research in Cognition: A Case Study of the AX-CPT. *Frontiers in Psychology*, *8*. https://www.frontiersin.org/articles/10.3389/fpsyg.2017.01482

Core Team, R. (2023). *R: A language and environment for statistical computing.* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Council, N. R., Medicine, I. of, Families, B. on C., Youth, and, & Development, C. on I. the S. of E. C. (2000). *From Neurons to Neighborhoods: The Science of Early Childhood Development*. National Academies Press.

Crespo-García, M., Zeiller, M., Leupold, C., Kreiselmeyer, G., Rampp, S., Hamer, H. M., & Dalal, S. S. (2016). Slow-theta power decreases during item-place encoding predict spatial accuracy of subsequent context recall. *NeuroImage*, *142*, 533–543. https://doi.org/10.1016/j.neuroimage.2016.08.021

Cross-Villasana, F., Finke, K., Hennig-Fast, K., Kilian, B., Wiegand, I., Müller, H. J., Möller, H.-J., & Töllner, T. (2015). The Speed of Visual Attention and Motor-Response Decisions in Adult Attention-Deficit/Hyperactivity Disorder. *Biological Psychiatry*, *78*(2), 107–115. https://doi.org/10.1016/j.biopsych.2015.01.016

Crouse, J. J., LaMonica, H. M., Song, Y. J. C., Boulton, K. A., Rohleder, C., DeMayo, M. M., Wilson, C. E., Loblay, V., Hindmarsh, G., Stratigos, T., Krausz, M., Foo, N., Teo, M., Hunter, A., Guastella, A. J., Banati, R. B., Troy, J., & Hickie, I. B. (2023). Designing an App for Parents and Caregivers to Promote Cognitive and Socioemotional Development and Well-being Among Children Aged 0 to 5 Years in Diverse Cultural Settings: Scientific Framework. *JMIR Pediatrics and Parenting*, *6*, e38921. https://doi.org/10.2196/38921

Cruz-Garza, J. G., Brantley, J. A., Nakagome, S., Kontson, K., Megjhani, M., Robleto, D., & Contreras-Vidal, J. L. (2017). Deployment of Mobile EEG Technology in an Art Museum Setting: Evaluation of Signal Quality and Usability. *Frontiers in Human Neuroscience*, *11*. https://doi.org/10.3389/fnhum.2017.00527

Cuevas, K., Raj, V., & Bell, M. A. (2012a). A frequency band analysis of two-year-olds' memory processes. *International Journal of Psychophysiology*, *83*(3), 315–322. https://doi.org/10.1016/j.ijpsycho.2011.11.009

Cuevas, K., Raj, V., & Bell, M. A. (2012b). A frequency band analysis of two-year-olds' memory processes. *International Journal of Psychophysiology*, *83*(3), 315–322. https://doi.org/10.1016/j.ijpsycho.2011.11.009

Cuevas, K., Raj, V., & Bell, M. A. (2012c). A frequency band analysis of two-year-olds' memory processes. *International Journal of Psychophysiology*, *83*(3), 315–322. https://doi.org/10.1016/j.ijpsycho.2011.11.009

Cullipher, S., Hansen, S. J. R., & VandenPlas, J. R. (2018). Eye Tracking as a Research Tool: An Introduction. In *Eye Tracking for the Chemistry Education Researcher* (Vol. 1292, pp. 1–9). American Chemical Society. https://doi.org/10.1021/bk-2018-1292.ch001

Dang, J., Xiao, S., Zhang, T., Liu, Y., Jiang, B., & Mao, L. (2016). When the poor excel: Poverty facilitates procedural learning. *Scandinavian Journal of Psychology*, *57*(4), 288–291. https://doi.org/10.1111/sjop.12292

D'Angiulli, A., Herdman, A., Stapells, D., & Hertzman, C. (2008). Children's Event-Related Potentials of Auditory Selective Attention Vary With Their Socioeconomic Status. *Neuropsychology*, *22*, 293–300. https://doi.org/10.1037/0894-4105.22.3.293

D'angiulli, A., Van Roon, P., Weinberg, J., Oberlander, T., Grunau, R., Hertzman, C., & Maggi, S. (2012). Frontal EEG/ERP correlates of attentional processes, cortisol and motivational states in adolescents from lower and higher socioeconomic status. *Frontiers in Human Neuroscience*, *6*. https://www.frontiersin.org/articles/10.3389/fnhum.2012.00306

Danks, D., & Davis, I. (2023). Causal inference in cognitive neuroscience. *WIREs Cognitive Science*, *14*(5), e1650. https://doi.org/10.1002/wcs.1650

Dawson, G., Webb, S. J., & McPartland, J. (2005). Understanding the Nature of Face Processing Impairment in Autism: Insights From Behavioral and Electrophysiological Studies. *Developmental Neuropsychology*, *27*, 403–424. https://doi.org/10.1207/s15326942dn2703_6

de Jong, M., Verhoeven, M., Hooge, I. T. C., & van Baar, A. L. (2016). Introduction of the Utrecht Tasks for Attention in Toddlers Using Eye Tracking (UTATE): A Pilot Study. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00669

de Souza Almeida, R., Faria-Jr, A., & Klein, R. M. (2021). On the origins and evolution of the Attention Network Tests. *Neuroscience & Biobehavioral Reviews*, *126*, 560–572. https://doi.org/10.1016/j.neubiorev.2021.02.028

Debnath, R., Buzzell, G. A., Morales, S., Bowers, M. E., Leach, S. C., & Fox, N. A. (2020). The Maryland analysis of developmental EEG (MADE) pipeline. *Psychophysiology*, *57*(6), e13580. https://doi.org/10.1111/psyp.13580

Debrouwere, S., & Rosseel, Y. (2023). *Why statistical adjustment for social and economic background fails and what to do about it*. PsyArXiv. https://doi.org/10.31234/osf.io/6m2ye

Del Bianco, T., Haartsen, R., Mason, L., Leno, V. C., Springer, C., Potter, M., Mackay, W., Smit, P., Plessis, C. D., Brink, L., Johnson, M. H., Murphy, D., Loth, E., Odendaal, H., & Jones, E. J. H. (2024). The importance of decomposing periodic and aperiodic EEG signals for assessment of brain function in a global context. *Developmental Psychobiology*, *66*(4), e22484. https://doi.org/10.1002/dev.22484

Del Bianco, T., Mason, L., Lai, M.-C., Loth, E., Tillmann, J., Charman, T., Hayward, H., Gleissl, T., Buitelaar, J. K., Murphy, D. G. M., Baron-Cohen, S., Bölte, S., Johnson, M. H., Jones, E. J. H., & Group, T. E.-A. L. (2022). Unique dynamic profiles of social attention in autistic females. *Journal of Child Psychology and Psychiatry*, *63*(12), 1602–1614. https://doi.org/10.1111/jcpp.13630

Demir, Ö. E., Prado, J., & Booth, J. R. (2015). Parental Socioeconomic Status and the Neural Basis of Arithmetic: Differential Relations to Verbal and Visuo-spatial Representations. *Developmental Science*, *18*(5), 799–814. https://doi.org/10.1111/desc.12268

Deng, W. (Sophia), & Sloutsky, V. M. (2016). Selective attention, diffused attention, and the development of categorization. *Cognitive Psychology*, *91*, 24–62. https://doi.org/10.1016/j.cogpsych.2016.09.002

Denton, K., Flanagan, K. D., & West, J. (2002). *Children's Reading and Mathematics Achievement in Kindergarten and First Grade*. National Center for Education Statistics, Office of Educational Research and Improvement, U.S. Department of Education.

Department for Education. (2021, September 3). *Early years foundation stage (EYFS) statutory framework*. GOV.UK. https://www.gov.uk/government/publications/early-years-foundation-stage-framework--2

Diamond, A. (2002). Normal Development of Prefrontal Cortex from Birth to Young Adulthood: Cognitive Functions, Anatomy, and Biochemistry. In D. T. Stuss & R. T. Knight (Eds.), *Principles of Frontal Lobe Function* (pp. 466–503). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195134971.003.0029

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, *64*(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Dixon-Woods, M., Tarrant, C., Jackson, C. J., Jones, D. R., & Kenyon, S. (2011). Providing the results of research to participants: A mixed-method study of the benefits and challenges of a consultative approach. *Clinical Trials*, *8*(3), 330–341. https://doi.org/10.1177/1740774511403514

Dotterer, A. M., Iruka, I. U., & Pungello, E. (2012). Parenting, Race, and Socioeconomic Status: Links to School Readiness. *Family Relations*, *61*(4), 657–670. https://doi.org/10.1111/j.1741-3729.2012.00716.x

Douglas Bates, Martin Maechler, Ben Bolker, & Steve Walker. (2015). Fitting Linear Mixed-Effects Models Using {lme4}. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i0

Downey, A. S., Busta, E. R., Mancher, M., Botkin, J. R., & & National Academies of Sciences, Engineering, and Medicine. (2018). Principles for the Return of Individual Research Results: Ethical and Societal Considerations—Returning Individual Research Results to Participants—NCBI Bookshelf. In *Returning Individual Research Results to Participants: Guidance for a New Research Paradigm*. National Academies Press. https://www.ncbi.nlm.nih.gov/books/NBK525079/

Draheim, C., Tsukahara, J. S., Martin, J., Mashburn, C., & Engle, R. W. (2019). *A toolbox approach to improving the measurement of attention control*. PsyArXiv. https://doi.org/10.31234/osf.io/q985d

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*(6), 1428–1446. https://doi.org/10.1037/0012-1649.43.6.1428

Duncan, G. J., & Magnuson, K. (2012). Socioeconomic status and cognitive functioning: Moving from correlation to causation: Socioeconomic status and cognitive functioning. *Wiley Interdisciplinary Reviews: Cognitive Science*, *3*(3), 377–386. https://doi.org/10.1002/wcs.1176

Duncan, G. J., Yeung, W. J., Brooks-Gunn, J., & Smith, J. R. (1998). How Much Does Childhood Poverty Affect the Life Chances of Children? *American Sociological Review*, *63*(3), 406–423. JSTOR. https://doi.org/10.2307/2657556

Ecton, W. G., Bennett, C. T., Nienhusser, H. K., Castillo-Montoya, M., & Dougherty, S. M. (2021). If You Fund Them, Will They Come? Implications From a PhD Fellowship Program on Racial/Ethnic Student Diversity. *AERA Open*, *7*, 23328584211040485. https://doi.org/10.1177/23328584211040485

Education in England: Annual Report 2018. (2018). Education Policy Institute. https://epi.org.uk/publications-and-research/annual-report-2018/

Eisma, J., Rawls, E., Long, S., Mach, R., & Lamm, C. (2021). Frontal midline theta differentiates separate cognitive control strategies while still generalizing the need for cognitive control. *Scientific Reports*, *11*(1), 14641. https://doi.org/10.1038/s41598-021-94162-z

Elliott, L., Bachman, H. J., Carvalho Pereira, J., Coulanges, L., Duong, S., Montue, T., Miller, P., Libertus, M., & Votruba-Drzal, E. (2022). Self-regulation in toddlers and the emergence of pre-academic disparities. *Infant Behavior and Development*, *69*, 101779. https://doi.org/10.1016/j.infbeh.2022.101779

Ellis, B. J., Abrams, L. S., Masten, A. S., Sternberg, R. J., Tottenham, N., & Frankenhuis, W. E. (2020). Hidden talents in harsh environments. *Development and Psychopathology*, 1–19. https://doi.org/10.1017/S0954579420000887

Ellis, B. J., Bianchi, J., Griskevicius, V., & Frankenhuis, W. E. (2017). Beyond Risk and Protective Factors: An Adaptation-Based Approach to Resilience. *Perspectives on Psychological Science*, *12*(4), 561–587. https://doi.org/10.1177/1745691617693054

Ellis, B. J., Sheridan, M. A., Belsky, J., & McLaughlin, K. A. (2022). Why and how does early adversity influence development? Toward an integrated model of dimensions of environmental experience. *Development and Psychopathology*, 1–25. https://doi.org/10.1017/S0954579421001838

Elsabbagh, M., Fernandes, J., Webb, S. J., Dawson, G., Charman, T., Johnson, M. H., & British Autism Study of Infant Siblings Team. (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., Volein, A., Holmboe, K., Tucker, L., Csibra, G., Baron-Cohen, S., Bolton, P., Charman, T., Baird, G., & Johnson, M. H. (2009). Visual orienting in the early broader autism phenotype: Disengagement and facilitation. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *50*(5), 637–642. https://doi.org/10.1111/j.1469-7610.2008.02051.x

Elsabbagh, M., Volein, A., Holmboe, K., Tucker, L., Csibra, G., Baron-Cohen, S., Bolton, P., Charman, T., Baird, G., & Johnson, M. H. (2009). Visual orienting in the early broader autism phenotype: Disengagement and facilitation. *Journal of Child Psychology and Psychiatry*, *50*(5), 637–642. https://doi.org/10.1111/j.1469-7610.2008.02051.x

Fan, J., McCandliss, B. D., Fossella, J., Flombaum, J. I., & Posner, M. I. (2005). The activation of attentional networks. *NeuroImage*, *26*(2), 471–479. https://doi.org/10.1016/j.neuroimage.2005.02.004

Farah, M. J. (2017a). The Neuroscience of Socioeconomic Status: Correlates, Causes, and Consequences. *Neuron*, *96*(1), 56–71. https://doi.org/10.1016/j.neuron.2017.08.034

Farah, M. J. (2017b). The Neuroscience of Socioeconomic Status: Correlates, Causes, and Consequences. *Neuron*, *96*(1), 56–71. https://doi.org/10.1016/j.neuron.2017.08.034

Farah, M. J. (2018). Socioeconomic status and the brain: Prospects for neuroscience-informed policy. *Nature Reviews Neuroscience*, *19*(7), 428–438. https://doi.org/10.1038/s41583-018-0023-2

Farroni, T., Johnson, M. H., Menon, E., Zulian, L., Faraguna, D., & Csibra, G. (2005). Newborns’ preference for face-relevant stimuli: Effects of contrast polarity. *Proceedings of the National Academy of Sciences*, *102*(47), 17245–17250. https://doi.org/10.1073/pnas.0502205102

Fearon, P. (2019). Editorial: Closing in on causal links between environmental exposures and human development using observational data – “confound those confounders!” *Journal of Child Psychology and Psychiatry*, *60*(2), 115–118. https://doi.org/10.1111/jcpp.13019

Fell, J., Ludowig, E., Staresina, B. P., Wagner, T., Kranz, T., Elger, C. E., & Axmacher, N. (2011). Medial Temporal Theta/Alpha Power Enhancement Precedes Successful Memory Encoding: Evidence Based on Intracranial EEG. *Journal of Neuroscience*, *31*(14), 5392–5397. https://doi.org/10.1523/JNEUROSCI.3668-10.2011

Ferguson, H. J., Brunsdon, V. E. A., & Bradford, E. E. F. (2021). The developmental trajectories of executive function from adolescence to old age. *Scientific Reports*, *11*(1), Article 1. https://doi.org/10.1038/s41598-020-80866-1

Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, *16*(2), 234–248. https://doi.org/10.1111/desc.12019

Fernández, T., Harmony, T., Rodríguez, M., Reyes, A., Marosi, E., & Bernal, J. (1993). Test-Retest Reliability of EEG Spectral Parameters During Cognitive Tasks: I Absolute and Relative Power. *International Journal of Neuroscience*, *68*(3–4), 255–261. https://doi.org/10.3109/00207459308994280

Finn, A. S., Minas, J. E., Leonard, J. A., Mackey, A. P., Salvatore, J., Goetz, C., West, M. R., Gabrieli, C. F. O., & Gabrieli, J. D. E. (2017). Functional brain organization of working memory in adolescents varies in relation to family income and academic achievement. *Developmental Science*, *20*(5), e12450. https://doi.org/10.1111/desc.12450

Finnigan, S., & Robertson, I. H. (2011). Resting EEG theta power correlates with cognitive performance in healthy older adults. *Psychophysiology*, *48*(8), 1083–1087. https://doi.org/10.1111/j.1469-8986.2010.01173.x

Fiske, A., & Holmboe, K. (2019). Neural substrates of early executive function development. *Developmental Review*, *52*, 42–62. https://doi.org/10.1016/j.dr.2019.100866

Fitzpatrick, C., McKinnon, R. D., Blair, C. B., & Willoughby, M. T. (2014). Do preschool executive function skills explain the school readiness gap between advantaged and disadvantaged children? *Learning and Instruction*, *30*, 25–31. https://doi.org/10.1016/j.learninstruc.2013.11.003

Fitzpatrick, C., & Pagani, L. S. (2012). Toddler working memory skills predict kindergarten school readiness. *Intelligence*, *40*(2), 205–212. https://doi.org/10.1016/j.intell.2011.11.007

Fox, J., & Weisberg, S. (2019). *An {R} Companion to Applied Regression* [Computer software]. Sage. https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Fox, S. E., Levitt, P., & Iii, C. A. N. (2010). How the Timing and Quality of Early Experiences Influence the Development of Brain Architecture. *Child Development*, *81*(1), 28–40. https://doi.org/10.1111/j.1467-8624.2009.01380.x

Foxe, J. J., Simpson, G. V., & Ahlfors, S. P. (1998). Parieto-occipital ~1 0Hz activity reflects anticipatory state of visual attention mechanisms. *NeuroReport*, *9*(17), 3929.

Foxe, J., & Snyder, A. (2011). The Role of Alpha-Band Brain Oscillations as a Sensory Suppression Mechanism during Selective Attention. *Frontiers in Psychology*, *2*. https://www.frontiersin.org/articles/10.3389/fpsyg.2011.00154

Frankenhuis, W. E., & de Weerth, C. (2013). Does Early-Life Exposure to Stress Shape or Impair Cognition? *Current Directions in Psychological Science*, *22*(5), 407–412. https://doi.org/10.1177/0963721413484324

Friedman, N. P., & Miyake, A. (2017). Unity and Diversity of Executive Functions: Individual Differences as a Window on Cognitive Structure. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *86*, 186–204. https://doi.org/10.1016/j.cortex.2016.04.023

Friedman, N. P., Miyake, A., Robinson, J. L., & Hewitt, J. K. (2011). Developmental Trajectories in Toddlers' Self-restraint Predict Individual Differences in Executive Functions 14 Years Later: A Behavioral Genetic Analysis. *Developmental Psychology*, *47*(5), 1410–1430. https://doi.org/10.1037/a0023750

Gabard-Durnam, L. J., Mendez Leal, A. S., Wilkinson, C. L., & Levin, A. R. (2018). The Harvard Automated Processing Pipeline for Electroencephalography (HAPPE): Standardized Processing Software for Developmental and High-Artifact Data. *Frontiers in Neuroscience*, *12*. https://www.frontiersin.org/articles/10.3389/fnins.2018.00097

Garcini, L. M., Arredondo, M. M., Berry, O., Church, J. A., Fryberg, S., Thomason, M. E., & McLaughlin, K. A. (2022). Increasing diversity in developmental cognitive neuroscience: A roadmap for increasing representation in pediatric neuroimaging research. *Developmental Cognitive Neuroscience*, *58*, 101167. https://doi.org/10.1016/j.dcn.2022.101167

Garrett, D. D., Samanez-Larkin, G. R., MacDonald, S. W. S., Lindenberger, U., McIntosh, A. R., & Grady, C. L. (2013). Moment-to-moment brain signal variability: A next frontier in human brain mapping? *Neuroscience & Biobehavioral Reviews*, *37*(4), 610–624. https://doi.org/10.1016/j.neubiorev.2013.02.015

Gartstein, M. A., & Rothbart, M. K. (2003). Studying infant temperament via the Revised Infant Behavior Questionnaire. *Infant Behavior and Development*, *26*(1), 64–86. https://doi.org/10.1016/S0163-6383(02)00169-8

Geyer, S., Hemström, Ö., Peter, R., & Vågerö, D. (2006). Education, income, and occupational class cannot be used interchangeably in social epidemiology. Empirical evidence against a common practice. *Journal of Epidemiology & Community Health*, *60*(9), 804–810. https://doi.org/10.1136/jech.2005.041319

Gilmore, J. H., Knickmeyer, R. C., & Gao, W. (2018). Imaging structural and functional brain development in early childhood. *Nature Reviews Neuroscience*, *19*(3), Article 3. https://doi.org/10.1038/nrn.2018.1

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

Gorjan, D., Gramann, K., Pauw, K. D., & Marusic, U. (2022). Removal of movement-induced EEG artifacts: Current state of the art and guidelines. *Journal of Neural Engineering*, *19*(1), 011004. https://doi.org/10.1088/1741-2552/ac542c

Govindan, R. B., Massaro, A., Vezina, G., Tsuchida, T., Cristante, C., & du Plessis, A. (2017). Does relative or absolute EEG power have prognostic value in HIE setting? *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *128*(1), 14–15. https://doi.org/10.1016/j.clinph.2016.10.094

Gräff, J., & Mansuy, I. M. (2008). Epigenetic codes in cognition and behaviour. *Behavioural Brain Research*, *192*(1), 70–87. https://doi.org/10.1016/j.bbr.2008.01.021

Gredebäck, G., Johnson, S., & von Hofsten, C. (2009). Eye Tracking in Infancy Research. *Developmental Neuropsychology*, *35*(1), 1–19. https://doi.org/10.1080/87565640903325758

Green, K. H., Van De Groep, I. H., Te Brinke, L. W., van der Cruijsen, R., van Rossenberg, F., & El Marroun, H. (2022). A perspective on enhancing representative samples in developmental human neuroscience: Connecting science to society. *Frontiers in Integrative Neuroscience*, *16*, 981657. https://doi.org/10.3389/fnint.2022.981657

Griffith, S. F., Hagan, M. B., Heymann, P., Heflin, B. H., & Bagner, D. M. (2020). Apps As Learning Tools: A Systematic Review. *Pediatrics*, *145*(1), e20191579. https://doi.org/10.1542/peds.2019-1579

Grunwald, M., Weiss, T., Krause, W., Beyer, L., Rost, R., Gutberlet, I., & Gertz, H.-J. (2001). Theta power in the EEG of humans during ongoing processing in a haptic object recognition task. *Cognitive Brain Research*, *11*(1), 33–37. https://doi.org/10.1016/S0926-6410(00)00061-6

Guderian, S., Schott, B. H., Richardson-Klavehn, A., & Düzel, E. (2009). Medial temporal theta state before an event predicts episodic encoding success in humans. *Proceedings of the National Academy of Sciences*, *106*(13), 5365–5370. https://doi.org/10.1073/pnas.0900289106

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*, *31*(5), e2323. 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*, 102314. https://doi.org/10.1016/j.nicl.2020.102314

Haartsen, R., Mason, L., Braithwaite, E., Del Bianco, T., Johnson, M. H., & Jones, E. J. H. (2021). Reliability of an automated gaze-controlled paradigm for capturing neural responses during visual and face processing in toddlerhood. *Developmental Psychobiology*, *63*(7), Article 7.

Haartsen, R., van der Velde, B., Jones, E. J. H., Johnson, M. H., & Kemner, C. (2020). Using multiple short epochs optimises the stability of infant EEG connectivity parameters. *Scientific Reports*, *10*(1), Article 1. https://doi.org/10.1038/s41598-020-68981-5

Hackman, D. A., & Farah, M. J. (2009). Socioeconomic status and the developing brain. *Trends in Cognitive Sciences*, *13*(2), 65–73. https://doi.org/10.1016/j.tics.2008.11.003

Hackman, D. A., Farah, M. J., & Meaney, M. J. (2010). Socioeconomic status and the brain: Mechanistic insights from human and animal research. *Nature Reviews Neuroscience*, *11*(9), 651–659. https://doi.org/10.1038/nrn2897

Hair, E., Halle, T., Terry-Humen, E., Lavelle, B., & Calkins, J. (2006a). Children's school readiness in the ECLS-K: Predictions to academic, health, and social outcomes in first grade. *Early Childhood Research Quarterly*, *21*(4), 431–454. https://doi.org/10.1016/j.ecresq.2006.09.005

Hair, E., Halle, T., Terry-Humen, E., Lavelle, B., & Calkins, J. (2006b). Children's school readiness in the ECLS-K: Predictions to academic, health, and social outcomes in first grade. *Early Childhood Research Quarterly*, *21*(4), 431–454. https://doi.org/10.1016/j.ecresq.2006.09.005

Halasz, J., & Kaufman, P. (2020). Social Structure and the Individual. In *A sociology experiment*.

Hampton Wray, A., Stevens, C., Pakulak, E., Isbell, E., Bell, T., & Helen Neville. (2017). Development of selective attention in preschool-age children from lower socioeconomic status backgrounds. *Developmental Cognitive Neuroscience*, *26*, 101–111. https://doi.org/10.1016/j.dcn.2017.06.006

Han, H.-R., Xu, A., Mendez, K. J. W., Okoye, S., Cudjoe, J., Bahouth, M., Reese, M., Bone, L., & Dennison-Himmelfarb, C. (2021). Exploring community engaged research experiences and preferences: A multi-level qualitative investigation. *Research Involvement and Engagement*, *7*(1), 19. https://doi.org/10.1186/s40900-021-00261-6

Harmony, T., Alvarez, A., Pascual, R., Ramos, A., Marosi, E., Díaz De León, A. E., Valdés, P., & Becker, J. (1988). EEG Maturation on Children with Different Economic and Psychosocial Characteristics. *International Journal of Neuroscience*, *41*(1–2), 103–113. https://doi.org/10.3109/00207458808985747

Harrell, F. E. Jr. (2023). *Hmisc: Harrell Miscellaneous* (Version 5.1-1) [Computer software]. https://cran.r-project.org/web/packages/Hmisc/index.html

Health and Social Care Committee. (2019). *First 1000 days of life—Health and Social Care Committee—House of Commons*. https://publications.parliament.uk/pa/cm201719/cmselect/cmhealth/1496/149603.htm#_idTextAnchor000

Heilicher, M., Crombie, K. M., & Cisler, J. M. (2022). Test-Retest Reliability of fMRI During an Emotion Processing Task: Investigating the Impact of Analytical

Approaches on ICC Values. *Frontiers in Neuroimaging*, *1*. https://www.frontiersin.org/articles/10.3389/fnimg.2022.859792

Hendry, A., Jones, E. J. H., Bedford, R., Gliga, T., Charman, T., Johnson, M. H., Baron-Cohen, S., Blasi, A., Bolton, P., Cheung, H. M. C., Davies, K., Elsabbagh, M., Fernandes, J., Gammer, I., Green, J., Guiraud, J., Lloyd-Fox, S., Liew, M., Maris, H., … the BASIS Team. (2018). Developmental change in look durations predicts later effortful control in toddlers at familial risk for ASD. *Journal of Neurodevelopmental Disorders*, *10*(1), 3. https://doi.org/10.1186/s11689-017-9219-4

Hendry, A., Jones, E. J. H., & Charman, T. (2016). Executive function in the first three years of life: Precursors, predictors and patterns. *Developmental Review*, *42*, 1–33. https://doi.org/10.1016/j.dr.2016.06.005

Henning, A., Spinath, F. M., & Aschersleben, G. (2011). The link between preschoolers' executive function and theory of mind and the role of epistemic states. *Journal of Experimental Child Psychology*, *108*(3), 513–531. https://doi.org/10.1016/j.jecp.2010.10.006

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Hervé, E., Mento, G., Desnous, B., & François, C. (2022). Challenges and new perspectives of developmental cognitive EEG studies. *NeuroImage*, *260*, 119508. https://doi.org/10.1016/j.neuroimage.2022.119508

High, P. C., & and the Committee on Early Childhood, A., and Dependent Care and Council on School Health. (2008). School Readiness. *Pediatrics*, *121*(4), e1008–e1015. https://doi.org/10.1542/peds.2008-0079

Hillard, B., El-Baz, A. S., Sears, L., Tasman, A., & Sokhadze, E. M. (2013). Neurofeedback training aimed to improve focused attention and alertness in children with ADHD: A study of relative power of EEG rhythms using custom-made software application. *Clinical EEG and Neuroscience*, *44*(3), 193–202. https://doi.org/10.1177/1550059412458262

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

Hofstadter, M., & Reznick, J. S. (1996). Response Modality Affects Human Infant Delayed-Response Performance. *Child Development*, *67*(2), 646–658. https://doi.org/10.1111/j.1467-8624.1996.tb01757.x

Hollingshead, A. (2011). *Four Factor Index of Social Status* (Vol. 8, pp. 21–52).

Holm, I. A. (2017). Pediatric Issues in Return of Results and Incidental Findings: Weighing Autonomy and Best Interests. *Genetic Testing and Molecular Biomarkers*, *21*(3), 155–158. https://doi.org/10.1089/gtmb.2016.0414

Holmboe, K., Pasco Fearon, R. M., Csibra, G., Tucker, L. A., & Johnson, M. H. (2008). Freeze-Frame: A new infant inhibition task and its relation to frontal cortex tasks during infancy and early childhood. *Journal of Experimental Child Psychology*, *100*(2), 89–114. https://doi.org/10.1016/j.jecp.2007.09.004

Hood, B., Cole-Davies, V., & Dias, M. (2003). Looking and search measures of object knowledge in preschool children. *Developmental Psychology*, *39*(1), 61–70.

Hood, B. M., & Atkinson, J. (1993). Disengaging visual attention in the infant and adult. *Infant Behavior and Development*, *16*(4), 405–422. https://doi.org/10.1016/0163-6383(93)80001-O

Huber, D. E., Potter, K. W., & Huszar, L. D. (2019). Less "Story" and more "Reliability" in cognitive neuroscience. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *113*, 347–349. https://doi.org/10.1016/j.cortex.2018.10.030

Humphreys, K. L., Lee, S. S., Telzer, E. H., Gabard-Durnam, L. J., Goff, B., Flannery, J., & Tottenham, N. (2015). Exploration—Exploitation strategy is dependent on early experience. *Developmental Psychobiology*, *57*(3), 313–321. https://doi.org/10.1002/dev.21293

Hutchinson, J. B., & Turk-Browne, N. B. (2012). Memory-guided attention: Control from multiple memory systems. *Trends in Cognitive Sciences*, *16*(12), 576–579. https://doi.org/10.1016/j.tics.2012.10.003

Jacobs, E. A., Rolle, I., Ferrans, C. E., Whitaker, E. E., & Warnecke, R. B. (2006). Understanding African Americans' views of the trustworthiness of physicians. *Journal of General Internal Medicine*, *21*(6), 642–647. https://doi.org/10.1111/j.1525-1497.2006.00485.x

Janus, M., & Duku, E. (2007). The School Entry Gap: Socioeconomic, Family, and Health Factors Associated With Children's School Readiness to Learn. *Early Education & Development*, *18*(3), 375–403. https://doi.org/10.1080/10409280701610796a

Jensen, S. K. G., Xie, W., Kumar, S., Haque, R., Petri, W. A., & Nelson, C. A. (2021). Associations of socioeconomic and other environmental factors with early brain development in Bangladeshi infants and children. *Developmental Cognitive Neuroscience*, *50*, 100981. https://doi.org/10.1016/j.dcn.2021.100981

Jiang, J., Bailey, K., & Xiao, X. (2018). Midfrontal Theta and Posterior Parietal Alpha Band Oscillations Support Conflict Resolution in a Masked Affective Priming Task. *Frontiers in Human Neuroscience*, *12*. https://doi.org/10.3389/fnhum.2018.00175

Jiang, Z., Waters, A. C., Liu, Y., Li, W., & Yang, L. (2017). Event-related theta oscillatory substrates for facilitation and interference effects of negative emotion on children's cognition. *International Journal of Psychophysiology*, *116*, 26–31. https://doi.org/10.1016/j.ijpsycho.2017.02.012

Johnson, M. H. (2001). Functional brain development in humans. *Nature Reviews Neuroscience*, *2*(7), Article 7. https://doi.org/10.1038/35081509

Jokisch, D., & Jensen, O. (2007). Modulation of Gamma and Alpha Activity during a Working Memory Task Engaging the Dorsal or Ventral Stream. *Journal of Neuroscience*, *27*(12), 3244–3251. https://doi.org/10.1523/JNEUROSCI.5399-06.2007

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), Article 1. https://doi.org/10.1038/s41598-020-67687-y

Jones, E. J. H., Mason, L., Begum Ali, J., van den Boomen, C., Braukmann, R., Cauvet, E., Demurie, E., Hessels, R. S., Ward, E. K., Hunnius, S., Bolte, S., Tomalski, P., Kemner, C., Warreyn, P., Roeyers, H., Buitelaar, J., Falck-Ytter, T., Charman, T., Johnson, M. H., & Eurosibs Team. (2019). Eurosibs: Towards robust measurement of infant neurocognitive predictors of autism across Europe. *Infant Behavior & Development*, *57*, 101316. https://doi.org/10.1016/j.infbeh.2019.03.007

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

Jordan, A., & Prendella, K. (2019). The invisible children of media research. *Journal of Children and Media*, *13*(2), 235–240. https://doi.org/10.1080/17482798.2019.1591662

Kaiser, J., Iliopoulos, P., Steinmassl, K., & Schütz-Bosbach, S. (2022). Preparing for Success: Neural Frontal Theta and Posterior Alpha Dynamics during Action Preparation Predict Flexible Resolution of Cognitive Conflicts. *Journal of Cognitive Neuroscience*, *34*(6), 1070–1089. https://doi.org/10.1162/jocn_a_01846

Kaldy, Z., Kraper, C., Carter, A. S., & Blaser, E. (2011). Toddlers with Autism Spectrum Disorder are more successful at visual search than typically developing toddlers. *Developmental Science*, *14*(5), 980–988. https://doi.org/10.1111/j.1467-7687.2011.01053.x

Kamil, K. (2017, February 19). Fifty People One Question- Galway, Ireland. *KamilFilms ○ Kamil Krolak*. https://www.kamilfilms.com/fifty-people-one-question-galway-ireland/

Karatekin, C. (2007). Eye tracking studies of normative and atypical development. *Developmental Review*, *27*(3), 283–348. https://doi.org/10.1016/j.dr.2007.06.006

Kassambara, A., & Mundt, F. (2020). *factoextra: Extract and Visualize the Results of Multivariate Data Analyses* (Version 1.0.7) [Computer software]. https://cran.r-project.org/web/packages/factoextra/index.html

Kennedy, B. R., Mathis, C. C., & Woods, A. K. (2007). African Americans and their distrust of the health care system: Healthcare for diverse populations. *Journal of Cultural Diversity*, *14*(2), 56–60.

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks Effect: Human Infants Allocate Attention to Visual Sequences That Are Neither Too Simple Nor Too Complex. *PLoS ONE*, *7*(5), e36399–e36399.

Kim, H.-Y. (2017). Statistical notes for clinical researchers: Chi-squared test and Fisher's exact test. *Restorative Dentistry & Endodontics*, *42*(2), 152–155. https://doi.org/10.5395/rde.2017.42.2.152

Kim, P., Capistrano, C., & Congleton, C. (2016). Socioeconomic disadvantages and neural sensitivity to infant cry: Role of maternal distress. *Social Cognitive and Affective Neuroscience*, *11*(10), 1597–1607. https://doi.org/10.1093/scan/nsw063

Kim, P., Evans, G. W., Angstadt, M., Ho, S. S., Sripada, C. S., Swain, J. E., Liberzon, I., & Phan, K. L. (2013). Effects of childhood poverty and chronic stress on emotion regulatory brain function in adulthood. *Proceedings of the National Academy of Sciences*, *110*(46), 18442–18447. https://doi.org/10.1073/pnas.1308240110

Kim, S.-P. (2018). Preprocessing of EEG. In C.-H. Im (Ed.), *Computational EEG Analysis: Methods and Applications* (pp. 15–33). Springer. https://doi.org/10.1007/978-981-13-0908-3_2

Kishiyama, M. M., Boyce, W. T., Jimenez, A. M., Perry, L. M., & Knight, R. T. (2009). Socioeconomic disparities affect prefrontal function in children. *Journal of Cognitive Neuroscience*, *21*(6), 1106–1115. https://doi.org/10.1162/jocn.2009.21101

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in psychtoolbox-3. *Perception*, *36*(14), 1–16.

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

Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences*, *16*(12), 606–617. https://doi.org/10.1016/j.tics.2012.10.007

Klimesch, W., Doppelmayr, M., Schwaiger, J., Auinger, P., & Winkler, Th. (1999). `Paradoxical' alpha synchronization in a memory task. *Cognitive Brain Research*, *7*(4), 493–501. https://doi.org/10.1016/S0926-6410(98)00056-1

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition–timing hypothesis. *Brain Research Reviews*, *53*(1), 63–88. https://doi.org/10.1016/j.brainresrev.2006.06.003

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 & Biobehavioral Reviews*, *50*, 189–203. https://doi.org/10.1016/j.neubiorev.2014.10.006

Knudsen, E. I., Heckman, J. J., Cameron, J. L., & Shonkoff, J. P. (2006). Economic, neurobiological, and behavioral perspectives on building America's future workforce. *Proceedings of the National Academy of Sciences*, *103*(27), 10155–10162. https://doi.org/10.1073/pnas.0600888103

Kolb, B., & Gibb, R. (2011). Brain Plasticity and Behaviour in the Developing Brain. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, *20*(4), 265–276.

Koller, M. (2016). robustlmm: An R Package for Robust Estimation of Linear Mixed-Effects Models. *Journal of Statistical Software*, *75*, 1–24. https://doi.org/10.18637/jss.v075.i06

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kopp, C. B. (1982). Antecedents of self-regulation: A developmental perspective. *Developmental Psychology*, *18*(2), 199–214. https://doi.org/10.1037/0012-1649.18.2.199

Kraybill, J. H., Kim-Spoon, J., & Bell, M. A. (2019). Infant Attention and Age 3 Executive Function. *The Yale Journal of Biology and Medicine*, *92*(1), 3–11.

Krishnan, S., & Johnson, M. H. (2014). A review of behavioural and brain development in the early years: The “toolkit” for later book-related skills. *Booktrust*.

Kucewicz, M. T., & Kamiński, J. (2022). Cognitive neuroscience: Theta network oscillations coordinate development of episodic memory. *Current Biology*, *32*(7), R331–R333. https://doi.org/10.1016/j.cub.2022.02.038

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(1), Article 1. https://doi.org/10.18637/jss.v082.i13

Lau-Zhu, A., Lau, M. P. H., & McLoughlin, G. (2019a). Mobile EEG in research on neurodevelopmental disorders: Opportunities and challenges. *Developmental Cognitive Neuroscience*, *36*, 100635. https://doi.org/10.1016/j.dcn.2019.100635

Lau-Zhu, A., Lau, M. P. H., & McLoughlin, G. (2019b). Mobile EEG in research on neurodevelopmental disorders: Opportunities and challenges. *Developmental Cognitive Neuroscience*, *36*, 100635. https://doi.org/10.1016/j.dcn.2019.100635

Lawson, G. M., Hook, C. J., & Farah, M. J. (2018). A meta-analysis of the relationship between socioeconomic status and executive function performance among children. *Developmental Science*, *21*(2), 10.1111/desc.12529. https://doi.org/10.1111/desc.12529

Leach, S. C., Morales, S., Bowers, M. E., Buzzell, G. A., Debnath, R., Beall, D., & Fox, N. A. (2020). Adjusting ADJUST: Optimizing the ADJUST algorithm for pediatric data using geodesic nets. *Psychophysiology*, *57*(8), e13566. https://doi.org/10.1111/psyp.13566

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

Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, *21*(1), 59–80. https://doi.org/10.1348/026151003321164627

Lenartowicz, A., Delorme, A., Walshaw, P. D., Cho, A. L., Bilder, R. M., McGough, J. J., McCracken, J. T., Makeig, S., & Loo, S. K. (2014). Electroencephalography Correlates of Spatial Working Memory Deficits in Attention-Deficit/Hyperactivity Disorder: Vigilance, Encoding, and Maintenance. *Journal of Neuroscience*, *34*(4), 1171–1182. https://doi.org/10.1523/JNEUROSCI.1765-13.2014

Lenth, R. V. (2023). *emmeans: Estimated Marginal Means, aka Least-Squares Means* [Computer software]. https://CRAN.R-project.org/package=emmeans

Leonard, J. A., Romeo, R. R., Park, A. T., Takada, M. E., Robinson, S. T., Grotzinger, H., Last, B. S., Finn, A. S., Gabrieli, J. D. E., & Mackey, A. P. (2019). Associations between cortical thickness and reasoning differ by socioeconomic status in development. *Developmental Cognitive Neuroscience*, *36*, 100641. https://doi.org/10.1016/j.dcn.2019.100641

Leone, L., Perugini, M., Bagozzi, R. P., Pierro, A., & Mannetti, L. (2001). Construct validity and generalizability of the Carver–White behavioural inhibition system/behavioural activation system scales. *European Journal of Personality*, *15*(5), 373–390. https://doi.org/10.1002/per.415

Leppänen, J. M. (2016). Using Eye Tracking to Understand Infants' Attentional Bias for Faces. *Child Development Perspectives*, *10*(3), 161–165. https://doi.org/10.1111/cdep.12180

Li-Grining, C. P., Votruba-Drzal, E., Maldonado-Carreño, C., & Haas, K. (2010a). Children's early approaches to learning and academic trajectories through fifth grade. *Developmental Psychology*, *46*(5), 1062–1077. https://doi.org/10.1037/a0020066

Li-Grining, C. P., Votruba-Drzal, E., Maldonado-Carreño, C., & Haas, K. (2010b). Children's early approaches to learning and academic trajectories through fifth grade. *Developmental Psychology*, *46*(5), 1062–1077. https://doi.org/10.1037/a0020066

Lindsley, D. B. (1939). A Longitudinal Study of the Occipital Alpha Rhythm in Normal Children: Frequency and Amplitude Standards. *The Pedagogical Seminary and Journal of Genetic Psychology*, *55*(1), 197–213. https://doi.org/10.1080/08856559.1939.10533190

Lipina, S., Martelli, M., & Colombo, J. (2005). Performance on the A-not-B task of Argentinean infants from unsatisfied and satisfied basic needs homes. *Revista Interamericana de Psicología = Interamerican Journal of Psychology, ISSN 0034-9690, Vol. 39, Nº. 1, 2005, Pags. 49-60*, *39*.

Lippe, S., Kovacevic, N., & McIntosh, R. (2009). Differential maturation of brain signal complexity in the human auditory and visual system. *Frontiers in Human Neuroscience*, *3*. https://www.frontiersin.org/articles/10.3389/neuro.09.048.2009

Lloyd-Fox, S., Papademetriou, M., Darboe, M. K., Everdell, N. L., Wegmuller, R., Prentice, A. M., Moore, S. E., & Elwell, C. E. (2014). Functional near infrared spectroscopy (fNIRS) to assess cognitive function in infants in rural Africa. *Scientific Reports*, *4*(1), Article 1. https://doi.org/10.1038/srep04740

Lockwood Estrin, G., Bhavnani, S., Goodwin, A., Arora, R., Divan, G., Haartsen, R., Mason, L., Patel, V., Johnson, M., & Jones, E. J. H. (2022). *From the lab to the field: Acceptability of using electroencephalography with Indian preschool childre*. https://doi.org/10.12688/wellcomeopenres.17334.1

Lockwood Estrin, G., Mason, L., Bhavnani, S., Gulati, S., Gliga, T., & Johnson, M. H. (2023). Attention control in autism: Eye-tracking findings from preschool children in a low- and middle-income country setting. *Autism*. https://doi.org/10.1177/13623613221149541

Luby, J., Belden, A., Botteron, K., Marrus, N., Harms, M. P., Babb, C., Nishino, T., & Barch, D. (2013). The Effects of Poverty on Childhood Brain Development: The

Mediating Effect of Caregiving and Stressful Life Events. *JAMA Pediatrics*, *167*(12), 1135–1142. https://doi.org/10.1001/jamapediatrics.2013.3139

Ma, M. K.-H., Fong, M. C.-M., & Wang, W. S. Y. (2019). Reliability of restingstate EEG spectral power—Advantage of normalization is not guaranteed. *26th Annual Meeting for the Cognitive Neuroscience Society. San Francisco, US*.

Maechler, M., original), P. R. (Fortran, original), A. S. (S, original), M. H. (S, Hornik [trl, K., maintenance(1999-2000)), ctb] (port to R., Studer, M., Roudier, P., Gonzalez, J., Kozlowski, K., pam()), E. S. (fastpam options for, & Murphy (volume.ellipsoid({d >= 3})), K. (2022). *cluster: 'Finding Groups in Data': Cluster Analysis Extended Rousseeuw et al.* (Version 2.1.4) [Computer software]. https://cran.r-project.org/web/packages/cluster/index.html

Maguire, M. J., & Schneider, J. M. (2019). Socioeconomic status related differences in resting state EEG activity correspond to differences in vocabulary and working memory in grade school. *Brain and Cognition*, *137*, 103619. https://doi.org/10.1016/j.bandc.2019.103619

Mann, C. A., Lubar, J. F., Zimmerman, A. W., Miller, C. A., & Muenchen, R. A. (1992). Quantitative analysis of EEG in boys with attention-deficit-hyperactivity disorder: Controlled study with clinical implications. *Pediatric Neurology*, *8*(1), 30–36. https://doi.org/10.1016/0887-8994(92)90049-5

Markant, J., Ackerman, L. K., Nussenbaum, K., & Amso, D. (2016). Selective attention neutralizes the adverse effects of low socioeconomic status on memory in 9-month-old infants. *Developmental Cognitive Neuroscience*, *18*, 26–33. https://doi.org/10.1016/j.dcn.2015.10.009

Markovska-Simoska, S., & Pop-Jordanova, N. (2017). Quantitative EEG in Children and Adults With Attention Deficit Hyperactivity Disorder: Comparison of Absolute and Relative Power Spectra and Theta/Beta Ratio. *Clinical EEG and Neuroscience*, *48*(1), 20–32. https://doi.org/10.1177/1550059416643824

Marshall, P. J., Bar-Haim, Y., & Fox, N. A. (2002). Development of the EEG from 5 months to 4 years of age. *Clinical Neurophysiology*, *113*(8), 1199–1208. https://doi.org/10.1016/S1388-2457(02)00163-3

Marshall, P. J., Fox, N. A., & Group, B. C. (2004). A Comparison of the Electroencephalogram between Institutionalized and Community Children in Romania. *Journal of Cognitive Neuroscience*, *16*(8), 1327–1338. https://doi.org/10.1162/0898929042304723

Marshall, P. J., Reeb, B. C., Fox, N. A., Nelson, C. A., & Zeanah, C. H. (2008). Effects of early intervention on EEG power and coherence in previously institutionalized children in Romania. *Development and Psychopathology*, *20*(3), 861–880. https://doi.org/10.1017/S0954579408000412

Marshall, P. J., Young, T., & Meltzoff, A. N. (2011). Neural correlates of action observation and execution in 14-month-old infants: An event-related EEG desynchronization study. *Developmental Science*, *14*(3), 474–480. https://doi.org/10.1111/j.1467-7687.2010.00991.x

Mashburn, A. J., & Pianta, R. C. (2006). Social Relationships and School Readiness. *Early Education & Development*, *17*(1), 151–176. https://doi.org/10.1207/s15566935eed1701_7

Matheny, A. P., Wachs, T. D., Ludwig, J. L., & Phillips, K. (1995). Bringing order out of chaos: Psychometric characteristics of the confusion, hubbub, and order scale. *Journal of Applied Developmental Psychology*, *16*(3), 429–444. https://doi.org/10.1016/0193-3973(95)90028-4

Mathewson, K. E., Kuziek, J. P., Scanlon, J. E. M., & Robles, D. (2024). The moving wave: Applications of the mobile EEG approach to study human attention. *Psychophysiology*, *61*(9), e14603. https://doi.org/10.1111/psyp.14603

Mathie, E., Wythe, H., Munday, D., Millac, P., Rhodes, G., Roberts, N., Smeeton, N., Poland, F., & Jones, J. (2018). Reciprocal relationships and the importance of feedback in patient and public involvement: A mixed methods study. *Health Expectations*, *21*(5), 899–908. https://doi.org/10.1111/hex.12684

Matsubara, K., Hattori, T., & Narumi, S. (2022). Achievement of Developmental Milestones Recorded in Real Time: A Mobile App-Based Study. *The Journal of Pediatrics*, *245*, 201-207.e9. https://doi.org/10.1016/j.jpeds.2022.02.018

Matte-Gagné, C., & Bernier, A. (2011). Prospective relations between maternal autonomy support and child executive functioning: Investigating the mediating role of child language ability. *Journal of Experimental Child Psychology*, *110*(4), 611–625. https://doi.org/10.1016/j.jecp.2011.06.006

Mayes, L. C., Carter, A. S., & Stubbe, D. (1993). Individual differences in exploratory behavior in the second year of life. *Infant Behavior and Development*, *16*(3), 269–284. https://doi.org/10.1016/0163-6383(93)80035-7

McClelland, M. M., Acock, A. C., & Morrison, F. J. (2006). The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school. *Early Childhood Research Quarterly*, *21*(4), 471–490. https://doi.org/10.1016/J.ECRESQ.2006.09.003

McCoy, D. C., Raver, C. C., & Sharkey, P. (2015). Children’s Cognitive Performance and Selective Attention Following Recent Community Violence. *Journal of Health and Social Behavior*, *56*(1), 19–36. https://doi.org/10.1177/0022146514567576

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, *1*(1), 30–46. https://doi.org/10.1037/1082-989X.1.1.30

McIntosh, A. R., Kovacevic, N., & Itier, R. J. (2008). Increased Brain Signal Variability Accompanies Lower Behavioral Variability in Development. *PLOS Computational Biology*, *4*(7), e1000106. https://doi.org/10.1371/journal.pcbi.1000106

McLaughlin, K. A., Fox, N. A., Zeanah, C. H., Sheridan, M. A., Marshall, P., & Nelson, C. A. (2010). Delayed Maturation in Brain Electrical Activity Partially Explains the Association Between Early Environmental Deprivation and Symptoms of Attention-Deficit/Hyperactivity Disorder. *Biological Psychiatry*, *68*(4), 329–336. https://doi.org/10.1016/j.biopsych.2010.04.005

McLaughlin, K. A., Sheridan, M. A., Humphreys, K. L., Belsky, J., & Ellis, B. J. (2021). The Value of Dimensional Models of Early Experience: Thinking Clearly about Concepts and Categories. *Perspectives on Psychological Science : A Journal of the Association for Psychological Science*, *16*(6), 1463. https://doi.org/10.1177/1745691621992346

McLaughlin, K. A., Weissman, D., & Bitrán, D. (2019). Childhood Adversity and Neural Development: A Systematic Review. *Annual Review of Developmental Psychology*, *1*(1), 277–312. https://doi.org/10.1146/annurev-devpsych-121318-084950

McWayne, C. M., Fantuzzo, J. W., & McDermott, P. A. (2004). Preschool Competency in Context: An Investigation of the Unique Contribution of Child Competencies to Early Academic Success. *Developmental Psychology*, *40*(4), 633–645. https://doi.org/10.1037/0012-1649.40.4.633

Meng, W., Black, C. F. D., & Feng, M. (2024). Associations among family socioeconomic status, parenting and sustained attention for socioeconomically disadvantaged children at age 5 years. *Infant and Child Development*, *33*(4), e2491. https://doi.org/10.1002/icd.2491

Merz, E. C., Wiltshire, C. A., & Noble, K. G. (2019). Socioeconomic Inequality and the Developing Brain: Spotlight on Language and Executive Function. *Child Development Perspectives*, *13*(1), 15–20. https://doi.org/10.1111/cdep.12305

Metzen, D., Genç, E., Getzmann, S., Larra, M. F., Wascher, E., & Ocklenburg, S. (2022). Frontal and parietal EEG alpha asymmetry: A large-scale investigation of short-term reliability on distinct EEG systems. *Brain Structure and Function*, *227*(2), 725–740. https://doi.org/10.1007/s00429-021-02399-1

Meyer, J. S. (1998). Chapter 22 - Behavioral Assessment in Developmental Neurotoxicology: Approaches Involving Unconditioned Behaviors Pharmacologic Challenges in Rodents. In W. Slikker & L. W. Chang (Eds.), *Handbook of Developmental Neurotoxicology* (pp. 403–426). Academic Press. https://doi.org/10.1016/B978-012648860-9.50029-7

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), Article 1. https://doi.org/10.1038/s41598-019-42615-x

Mezzacappa, E. (2004). Alerting, Orienting, and Executive Attention: Developmental Properties and Sociodemographic Correlates in an Epidemiological Sample of Young, Urban Children. *Child Development*, *75*(5), 1373–1386. https://doi.org/10.1111/j.1467-8624.2004.00746.x

Micalizzi, L., Brick, L. A., Flom, M., Ganiban, J. M., & Saudino, K. J. (2019). Effects of socioeconomic status and executive function on school readiness across levels of household chaos. *Early Childhood Research Quarterly*, *47*, 331–340. https://doi.org/10.1016/j.ecresq.2019.01.007

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202. https://doi.org/10.1146/annurev.neuro.24.1.167

Miriti, M. N. (2020). The Elephant in the Room: Race and STEM Diversity. *BioScience*, *70*(3), 237–242. https://doi.org/10.1093/biosci/biz167

Mittal, C., Griskevicius, V., Simpson, J. A., Sung, S., & Young, E. S. (2015). Cognitive adaptations to stressful environments: When childhood adversity enhances adult executive function. *Journal of Personality and Social Psychology*, *109*, 604–621. https://doi.org/10.1037/pspi0000028

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex “Frontal Lobe” Tasks: A Latent Variable Analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Morales, M., Mundy, P., Crowson, M., Neal, A. R., & Delgado, C. (2005). Individual differences in infant attention skills, joint attention, and emotion regulation behaviour. *International Journal of Behavioral Development*, *29*(3), 259–263. https://doi.org/10.1080/01650250444000432

Morales, S., & Bowers, M. E. (2022). Time-frequency analysis methods and their application in developmental EEG data. *Developmental Cognitive Neuroscience*, *54*, 101067. https://doi.org/10.1016/j.dcn.2022.101067

Moriguchi, Y., & Shinohara, I. (2019). Socioeconomic disparity in prefrontal development during early childhood. *Scientific Reports*, *9*(1), 2585. https://doi.org/10.1038/s41598-019-39255-6

Mujahid, M. S., Diez Roux, A. V., Morenoff, J. D., & Raghunathan, T. (2007). Assessing the Measurement Properties of Neighborhood Scales: From Psychometrics to Ecometrics. *American Journal of Epidemiology*, *165*(8), 858–867. https://doi.org/10.1093/aje/kwm040

Mulder, H., Hoofs, H., Verhagen, J., van der Veen, I., & Leseman, P. P. M. (2014). Psychometric properties and convergent and predictive validity of an executive function test battery for two-year-olds. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00733

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

Munsters, N. M., van Ravenswaaij, H., van den Boomen, C., & Kemner, C. (2019). Test-retest reliability of infant event related potentials evoked by faces. *Neuropsychologia*, *126*, 20–26. https://doi.org/10.1016/j.neuropsychologia.2017.03.030

National Statistics. (2019). *English indices of deprivation 2019*. GOV.UK. https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019

Nelson, C. A., & Gabard-Durnam, L. J. (2020). Early Adversity and Critical Periods: Neurodevelopmental Consequences of Violating the Expectable Environment. *Trends in Neurosciences*, *43*(3), 133–143. https://doi.org/10.1016/j.tins.2020.01.002

Nelson, C. A., Zeanah, C. H., & Fox, N. A. (2019). How Early Experience Shapes Human Development: The Case of Psychosocial Deprivation. *Neural Plasticity*, *2019*, e1676285. https://doi.org/10.1155/2019/1676285

Nelson, D. R., Adger, W. N., & Brown, K. (2007). Adaptation to Environmental Change: Contributions of a Resilience Framework. *Annual Review of Environment and Resources*, *32*(1), 395–419. https://doi.org/10.1146/annurev.energy.32.051807.090348

Nielsen, M., Haun, D., Kärtner, J., & Legare, C. H. (2017). The persistent sampling bias in developmental psychology: A call to action. *Journal of Experimental Child Psychology*, *162*, 31–38. https://doi.org/10.1016/j.jecp.2017.04.017

Nishiyori, R., Xiao, R., Vanderbilt, D., & Smith, B. A. (2021). Electroencephalography measures of relative power and coherence as reaching skill emerges in infants born preterm. *Scientific Reports*, *11*(1), Article 1. https://doi.org/10.1038/s41598-021-82329-7

Nketia, J., Amso, D., & Brito, N. H. (2021). Towards a more inclusive and equitable developmental cognitive neuroscience. *Developmental Cognitive Neuroscience*, *52*, 101014. https://doi.org/10.1016/j.dcn.2021.101014

Noble, K. G., Engelhardt, L. E., Brito, N. H., Mack, L. J., Nail, E. J., Angal, J., Barr, R., Fifer, W. P., Elliott, A. J., & Network, in collaboration with the P. (2015). Socioeconomic disparities in neurocognitive development in the first two years of life. *Developmental Psychobiology*, *57*(5), 535–551. https://doi.org/10.1002/dev.21303

Noble, K. G., Farah, M. J., & McCandliss, B. D. (2006). Socioeconomic background modulates cognition–achievement relationships in reading. *Cognitive Development*, *21*(3), 349–368. https://doi.org/10.1016/j.cogdev.2006.01.007

Noble, K. G., McCandliss, B. D., & Farah, M. J. (2007). Socioeconomic gradients predict individual differences in neurocognitive abilities. *Developmental Science*, *10*(4), 464–480. https://doi.org/10.1111/j.1467-7687.2007.00600.x

Nowakowski, S. (2023). *hclust1d: Hierarchical Clustering of Univariate (1d) Data* (Version 0.1.1) [Computer software]. https://cran.r-project.org/web/packages/hclust1d/index.html

Nyström, P., Ljunghammar, T., Rosander, K., & von Hofsten, C. (2011). Using mu rhythm desynchronization to measure mirror neuron activity in infants.

Developmental Science, *14*(2), 327–335. https://doi.org/10.1111/j.1467-7687.2010.00979.x

Office for National Statistics. (2022). *Ethnic group—Census Maps, ONS*. https://www.ons.gov.uk/census/maps/choropleth/identity/ethnic-group/ethnic-group-tb-20b/asian-asian-british-or-asian-welsh-bangladeshi

Office for National Statistics. (2023a). *Effects of taxes and benefits on UK household income—Office for National Statistics*. https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/bulletins/theeffectsoftaxesandbenefitsonhouseholdincome/financialyearending2022

Office for National Statistics. (2023b). *How qualification levels across England and Wales differ by country of birth—Office for National Statistics*. https://www.ons.gov.uk/peoplepopulationandcommunity/educationandchildcare/articles/howqualificationlevelsacrossenglandandwalesdifferbycountryofbirth/2023-05-15

Oliveira, A. S., Schlink, B. R., Hairston, W. D., König, P., & Ferris, D. P. (2016). Proposing Metrics for Benchmarking Novel EEG Technologies Towards Real-World Measurements. *Frontiers in Human Neuroscience*, *10*. https://doi.org/10.3389/fnhum.2016.00188

Olson, L., Chen, B., & Fishman, I. (2021). Neural Correlates of Socioeconomic Status in Early Childhood: A Systematic Review of the Literature. *Child Neuropsychology : A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *27*(3), 390–423. https://doi.org/10.1080/09297049.2021.1879766

Omand, J. A., Janus, M., Maguire, J. L., Parkin, P. C., Aglipay, M., Simpson, J. R., Keown-Stoneman, C. D. G., Duku, E., Reid-Westoby, C., & Birken, C. S. (2021). Nutritional Risk in Early Childhood and School Readiness. *The Journal of Nutrition*, *151*(12), 3811–3819. https://doi.org/10.1093/jn/nxab307

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, *2011*, 1:1-1:9. https://doi.org/10.1155/2011/156869

Orekhova, E. 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, E. 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

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

Otero, G. A. (1994). Eeg spectral analysis in children with sociocultural handicaps. *International Journal of Neuroscience*, *79*(3–4), 213–220. https://doi.org/10.3109/00207459408986082

Otero, G. A. (1997). Poverty, cultural disadvantage and brain development: A study of pre-school children in Mexico. *Electroencephalography and Clinical Neurophysiology*, *102*(6), 512–516. https://doi.org/10.1016/S0013-4694(97)95213-9

Otero, G. A., Pliego-Rivero, F. B., Fernández, T., & Ricardo, J. (2003). EEG development in children with sociocultural disadvantages: A follow-up study. *Clinical Neurophysiology*, *114*(10), 1918–1925. https://doi.org/10.1016/S1388-2457(03)00173-1

Oudgenoeg-Paz, O., Leseman, P. P. M., & Volman, M. (Chiel) J. M. (2014). Can infant self-locomotion and spatial exploration predict spatial memory at school age? *European Journal of Developmental Psychology*, *11*(1), 36–48. https://doi.org/10.1080/17405629.2013.803470

Pakulak, E., Stevens, C., & Neville, H. (2018). Neuro-, Cardio-, and Immunoplasticity: Effects of Early Adversity. *Annual Review of Psychology*, *69*(Volume 69, 2018), 131–156. https://doi.org/10.1146/annurev-psych-010416-044115

Panov, G. (2023). Quantitative EEG Analysis: Introduction and Basic Principles. In D. Stoyanov, B. Draganski, P. Brambilla, & C. Lamm (Eds.), *Computational Neuroscience* (pp. 85–91). Springer US. https://doi.org/10.1007/978-1-0716-3230-7_5

Passaretta, G., & Skopek, J. (2018). Roots and Development of Achievement Gaps. *Dublin: Trinity College.*, *ISOTIS Report (D 1.3)*.

Passaretta, G., Skopek, J., & Huizen, T. (2020). *To what extent is social inequality in school-age achievement determined before and during schooling? A longitudinal analysis in three European countries*. https://doi.org/10.31235/osf.io/yqt6n

Pavlov, Y. G., & Kotchoubey, B. (2022). Oscillatory brain activity and maintenance of verbal and visual working memory: A systematic review. *Psychophysiology*, *59*(5), e13735. https://doi.org/10.1111/psyp.13735

Peelen, M. V., Heslenfeld, D. J., & Theeuwes, J. (2004). Endogenous and exogenous attention shifts are mediated by the same large-scale neural network. *NeuroImage*, *22*(2), 822–830. https://doi.org/10.1016/j.neuroimage.2004.01.044

Pellegrini, A. D., & Smith, P. K. (2001). Play and Development in Children. In N. J. Smelser & P. B. Baltes (Eds.), *International Encyclopedia of the Social & Behavioral Sciences* (pp. 11501–11503). Pergamon. https://doi.org/10.1016/B0-08-043076-7/01676-4

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

Phillips, J. C., Alfano, A. R., Barfield, L. C., Cain, L., Sadjadi, M., Morales, E., Phillips-Beck, W., Galarza, M. G., Torres, M., Zindani, S., Rayani, A., Edwards, K., Jones, S. G., & Hannan, J. (2024). Exploring Maternal and Infant Health App Development and Effectiveness Research: Scoping Review. *JMIR Pediatrics and Parenting*, *7*. https://doi.org/10.2196/46973

Piñeyro Salvidegoitia, M., Jacobsen, N., Bauer, A.-K. R., Griffiths, B., Hanslmayr, S., & Debener, S. (2019). Out and about: Subsequent memory effect captured in a natural outdoor environment with smartphone EEG. *Psychophysiology*, *56*(5), e13331. https://doi.org/10.1111/psyp.13331

Popov, T., Tröndle, M., Baranczuk-Turska, Z., Pfeiffer, C., Haufe, S., & Langer, N. (2023). Test–retest reliability of resting-state EEG in young and older adults. *Psychophysiology*, *60*(7), e14268. https://doi.org/10.1111/psyp.14268

Portugal, A. M., Bedford, R., Cheung, C. H. M., Mason, L., & Smith, T. J. (2021). Longitudinal touchscreen use across early development is associated with faster exogenous and reduced endogenous attention control. *Scientific Reports*, *11*(1), Article 1. https://doi.org/10.1038/s41598-021-81775-7

Posner, M. I., & Petersen, S. E. (1990). The Attention System of the Human Brain. *Annual Review of Neuroscience*, *13*(1), 25–42. https://doi.org/10.1146/annurev.ne.13.030190.000325

Posner, M. I., Rothbart, M. K., Sheese, B. E., & Voelker, P. (2014). Developing Attention: Behavioral and Brain Mechanisms. *Advances in Neuroscience*, *2014*(1), 405094. https://doi.org/10.1155/2014/405094

Posner, M. I., Walker, J. A., Friedrich, F. A., & Rafal, R. D. (1987). How do the parietal lobes direct covert attention? *Neuropsychologia*, *25*(1), 135–145. https://doi.org/10.1016/0028-3932(87)90049-2

Putnam, S. P., Helbig, A. L., Gartstein, M. A., Rothbart, M. K., & Leerkes, E. (2014). Development and Assessment of Short and Very Short Forms of the Infant Behavior Questionnaire–Revised. *Journal of Personality Assessment*, *96*(4), 445–458. https://doi.org/10.1080/00223891.2013.841171

Raizada, R. D. S., Richards, T. L., Meltzoff, A., & Kuhl, P. K. (2008). Socioeconomic status predicts hemispheric specialisation of the left inferior frontal gyrus in young children. *NeuroImage*, *40*(3), 1392–1401. https://doi.org/10.1016/j.neuroimage.2008.01.021

Raver, C. C., Blair, C., & Willoughby, M. (2013). Poverty as a Predictor of 4-Year-Olds' Executive Function: New Perspectives on Models of Differential Susceptibility. *Developmental Psychology*, *49*(2), 292–304. https://doi.org/10.1037/a0028343

Ravitsky, V., & Wilfond, B. S. (2006). Disclosing Individual Genetic Results to Research Participants. *The American Journal of Bioethics*, *6*(6), 8–17. https://doi.org/10.1080/15265160600934772

Reikerås, E., & Dahle, A. E. (2022). Relations between reading skill level in fifth grade and functional language skills at toddler age. *European Early Childhood Education Research Journal*, *30*(6), 961–974. https://doi.org/10.1080/1350293X.2022.2046831

Reilly, E. B., Stallworthy, I. C., Mliner, S. B., Troy, M. F., Elison, J. T., & Gunnar, M. R. (2021). Infants' abilities to respond to cues for joint attention vary by family socioeconomic status. *Infancy : The Official Journal of the International Society on Infant Studies*, *26*(2), 204–222. https://doi.org/10.1111/infa.12380

Richmond, J., & Nelson, C. A. (2009). Relational memory during infancy: Evidence from eye tracking. *Developmental Science*, *12*(4), 549–556. https://doi.org/10.1111/j.1467-7687.2009.00795.x

Rogers, J. M., Johnstone, S. J., Aminov, A., Donnelly, J., & Wilson, P. H. (2016a). Test-retest reliability of a single-channel, wireless EEG system. *International Journal of Psychophysiology*, *106*, 87–96. https://doi.org/10.1016/j.ijpsycho.2016.06.006

Rogers, J. M., Johnstone, S. J., Aminov, A., Donnelly, J., & Wilson, P. H. (2016b). Test-retest reliability of a single-channel, wireless EEG system. *International Journal of Psychophysiology*, *106*, 87–96. https://doi.org/10.1016/j.ijpsycho.2016.06.006

Romeo, R. R., Flournoy, J. C., McLaughlin, K. A., & Lengua, L. J. (2022). Language development as a mechanism linking socioeconomic status to executive functioning development in preschool. *Developmental Science*, *25*(5). https://doi.org/10.1111/desc.13227

Rong, T., Ristevski, E., & Carroll, M. (2023). Exploring community engagement in place-based approaches in areas of poor health and disadvantage: A scoping review. *Health & Place*, *81*, 103026. https://doi.org/10.1016/j.healthplace.2023.103026

Rose, S., Feldman, J., & Jankowski, J. (2001). Attention and recognition memory in the 1st year of life: A longitudinal study of preterm and full-term infants. *Developmental Psychology*, *37*, 135–151. https://doi.org/10.1037//0012-1649.37.1.135

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*, 1–36. https://doi.org/10.18637/jss.v048.i02

Rowe, M. L., Raudenbush, S. W., & Goldin-Meadow, S. (2012). The pace of vocabulary growth helps predict later vocabulary skill. *Child Development*, *83*(2), 508–525. https://doi.org/10.1111/j.1467-8624.2011.01710.x

Rowley, S. J., & Camacho, T. C. (2015). Increasing Diversity in Cognitive Developmental Research: Issues and Solutions. *Journal of Cognition and Development*, *16*(5), 683–692. https://doi.org/10.1080/15248372.2014.976224

Ruiz, M. A., Zamorano, E., García-Campayo, J., Pardo, A., Freire, O., & Rejas, J. (2011). Validity of the GAD-7 scale as an outcome measure of disability in patients with generalized anxiety disorders in primary care. *Journal of Affective Disorders*, *128*(3), 277–286. https://doi.org/10.1016/j.jad.2010.07.010

Rusher, A. S., Cross, D. R., & Ware, A. M. (1995). Infant and toddler play: Assessment of exploratory style and development level. *Early Childhood Research Quarterly*, *10*(3), 297–315. https://doi.org/10.1016/0885-2006(95)90009-8

Saby, J. N., & Marshall, P. J. (2012). The Utility of EEG Band Power Analysis in the Study of Infancy and Early Childhood. *Developmental Neuropsychology*, *37*(3), 253–273. https://doi.org/10.1080/87565641.2011.614663

Saez de Urabain, I. R., Nuthmann, A., Johnson, M. H., & Smith, T. J. (2017). Disentangling the mechanisms underlying infant fixation durations in scene perception: A computational account. *Vision Research*, *134*, 43–59. https://doi.org/10.1016/j.visres.2016.10.015

Salarian, A. (2023, August 3). *Intraclass Correlation Coefficient (ICC)*. https://uk.mathworks.com/matlabcentral/fileexchange/22099-intraclass-correlation-coefficient-icc

Sarter, M., Givens, B., & Bruno, J. P. (2001). The cognitive neuroscience of sustained attention: Where top-down meets bottom-up. *Brain Research Reviews*, *35*(2), 146–160. https://doi.org/10.1016/S0165-0173(01)00044-3

Sasson, N. J., & Elison, J. T. (2012). Eye Tracking Young Children with Autism. *JoVE (Journal of Visualized Experiments)*, *61*, e3675. https://doi.org/10.3791/3675

Schliemann, A. D., & Carraher, D. W. (2002). The Evolution of Mathematical Reasoning: Everyday versus Idealized Understandings. *Developmental Review*, *22*(2), 242–266. https://doi.org/10.1006/drev.2002.0547

Sederberg, P. B., Kahana, M. J., Howard, M. W., Donner, E. J., & Madsen, J. R. (2003). Theta and Gamma Oscillations during Encoding Predict Subsequent Recall. *Journal of Neuroscience*, *23*(34), 10809–10814. https://doi.org/10.1523/JNEUROSCI.23-34-10809.2003

Segalowitz, S., Santesso, D., & Jetha, M. (2010). *Electrophysiological changes during adolescence: A review—ScienceDirect*. https://www.sciencedirect.com/science/article/pii/S027826260900195X

Senoussi, M., Verbeke, P., Desender, K., De Loof, E., Talsma, D., & Verguts, T. (2022). Theta oscillations shift towards optimal frequency for cognitive control. *Nature Human Behaviour*, *6*(7), 1000–1013. https://doi.org/10.1038/s41562-022-01335-5

Sheehy-Skeffington, J. (2020). The effects of low socioeconomic status on decision-making processes. *Current Opinion in Psychology*, *33*, 183–188. https://doi.org/10.1016/j.copsyc.2019.07.043

Sheese, B. E., Voelker, P., Posner, M. I., & Rothbart, M. K. (2009). Genetic variation influences on the early development of reactive emotions and their regulation by attention. *Cognitive Neuropsychiatry*, *14*(4–5), 332–355. https://doi.org/10.1080/13546800902844064

Sheridan, M. A., & McLaughlin, K. A. (2014). Dimensions of early experience and neural development: Deprivation and threat. *Trends in Cognitive Sciences*, *18*(11), 580–585. https://doi.org/10.1016/j.tics.2014.09.001

Sheridan, M. A., Sarsour, K., Jutte, D., D'Esposito, M., & Boyce, W. T. (2012). The Impact of Social Disparity on Prefrontal Function in Childhood. *PLOS ONE*, *7*(4), e35744. https://doi.org/10.1371/journal.pone.0035744

Shonkoff, J. P., Richter, L., Van Der Gaag, J., & Bhutta, Z. A. (2012). An Integrated Scientific Framework for Child Survival and Early Childhood Development. *Pediatrics*, *129*(2), e460–e472. https://doi.org/10.1542/peds.2011-0366

Shultz, S., Klin, A., & Jones, W. (2018). Neonatal Transitions in Social Behavior and Their Implications for Autism. *Trends in Cognitive Sciences*, *22*(5), 452–469. https://doi.org/10.1016/j.tics.2018.02.012

Siqueiros Sanchez, M., Ronald, A., Mason, L., Jones, E. J. H., Bölte, S., & Falck-Ytter, T. (2021). Visual disengagement in young infants in relation to age, sex, SES, developmental level and adaptive functioning. *Infant Behavior and Development*, *63*, 101555. https://doi.org/10.1016/j.infbeh.2021.101555

Skinner, J. S., Williams, N. A., Richmond, A., Brown, J., Strelnick, A. H., Calhoun, K., De Loney, E. H., Allen, S., Pirie, A., & Wilkins, C. H. (2018). Community Experiences and Perceptions of Clinical and Translational Research and Researchers. *Progress in Community Health Partnerships : Research, Education, and Action*, *12*(3), 263–271. https://doi.org/10.1353/cpr.2018.0050

Smith, J. R. (1939). The "Occipital" and "Pre-Central" Alpha Rhythms During the First two Years. *The Journal of Psychology*, *7*(2), 223–226. https://doi.org/10.1080/00223980.1939.9917630

Smith, J. R. (1941). The Frequency Growth of the Human Alpha Rhythms During Normal Infancy and Childhood. *The Journal of Psychology*, *11*(1), 177–198. https://doi.org/10.1080/00223980.1941.9917028

Snow, K. L. (2006). Measuring School Readiness: Conceptual and Practical Considerations. *Early Education and Development*, *17*(1), 7–41. https://doi.org/10.1207/s15566935eed1701_2

Somsen, R. J. M., Van'T Klooster, B. J., Van Der Molen, M. W., Van Leeuwen, H. M. P., & Licht, R. (1997). Growth spurts in brain maturation during middle childhood as indexed by EEG power spectra. *Biological Psychology*, *44*(3), 187–209. https://doi.org/10.1016/S0301-0511(96)05218-0

Southgate, V., Johnson, M. H., Osborne, T., & Csibra, G. (2009a). Predictive motor activation during action observation in human infants. *Biology Letters*, *5*(6), 769–772. https://doi.org/10.1098/rsbl.2009.0474

Southgate, V., Johnson, M. H., Osborne, T., & Csibra, G. (2009b). Predictive motor activation during action observation in human infants. *Biology Letters*, *5*(6), 769–772. https://doi.org/10.1098/rsbl.2009.0474

Spann, M. N., Bansal, R., Hao, X., Rosen, T. S., & Peterson, B. S. (2020). Prenatal Socioeconomic Status and Social Support are associated with Neonatal Brain Morphology, Toddler Language and Psychiatric Symptoms. *Child Neuropsychology : A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *26*(2), 170–188. https://doi.org/10.1080/09297049.2019.1648641

Sparrow, S. S., & Cicchetti, D. V. (1989). The Vineland Adaptive Behavior Scales. In *Major psychological assessment instruments, Vol. 2.* (pp. 199–231). Allyn & Bacon.

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7. *Archives of Internal Medicine*, *166*(10), 1092–1097. https://doi.org/10.1001/archinte.166.10.1092

St. John, A. M., Finch, K., & Tarullo, A. R. (2019). Socioeconomic status and neural processing of a go/no-go task in preschoolers: An assessment of the P3b. *Developmental Cognitive Neuroscience*, *38*, 100677. https://doi.org/10.1016/j.dcn.2019.100677

Stevens, C., & Bavelier, D. (2012). The role of selective attention on academic foundations: A cognitive neuroscience perspective. *Developmental Cognitive Neuroscience*, *2*, S30–S48. https://doi.org/10.1016/j.dcn.2011.11.001

Stevens, C., Lauinger, B., & Neville, H. (2009). Differences in the neural mechanisms of selective attention in children from different socioeconomic backgrounds: An event-related brain potential study. *Developmental Science*, *12*(4), 634–646. https://doi.org/10.1111/j.1467-7687.2009.00807.x

Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods*, *42*(4), 1096–1104. https://doi.org/10.3758/BRM.42.4.1096

Stoet, G. (2017). PsyToolkit: A Novel Web-Based Method for Running Online Questionnaires and Reaction-Time Experiments. *Teaching of Psychology*, *44*(1), 24–31. https://doi.org/10.1177/0098628316677643

Stroganova, T. A., & Orekhova, E. V. (2007). EEG and infant states. In *Infant EEG and Event-Related Potentials* (pp. 251–287). Psychology Press.

Stroganova, T. A., Orekhova, E. V., & Posikera, I. N. (1999). EEG alpha rhythm in infants. *Clinical Neurophysiology*, *110*(6), 997–1012. https://doi.org/10.1016/S1388-2457(98)00009-1

Swartz, T. H., Palermo, A.-G. S., Masur, S. K., & Aberg, J. A. (2019). The Science and Value of Diversity: Closing the Gaps in Our Understanding of Inclusion and Diversity. *The Journal of Infectious Diseases*, *220*(Supplement_2), S33–S41. https://doi.org/10.1093/infdis/jiz174

Szepsenwol, O., & Simpson, J. A. (2019). Attachment within life history theory: An evolutionary perspective on individual differences in attachment. *Current Opinion in Psychology*, *25*, 65–70. https://doi.org/10.1016/j.copsyc.2018.03.005

Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (Seventh edition). Pearson.

Throm, E., Gui, A., Haartsen, R., da Costa, P. F., Leech, R., & Jones, E. J. H. (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

Tierney, A. L., & Nelson, C. A. (2009). Brain Development and the Role of Experience in the Early Years. *Zero to Three*, *30*(2), 9–13.

Tomalski, P., Moore, D. G., Ribeiro, H., Axelsson, E. L., Murphy, E., Karmiloff-Smith, A., Johnson, M. H., & Kushnerenko, E. (2013). Socioeconomic status and functional brain development—Associations in early infancy. *Developmental Science*, *16*(5), 676–687. https://doi.org/10.1111/desc.12079

Troller-Renfree, S., Morales, S., Leach, S., Bowers, M., Debnath, R., Fifer, W., Fox, N., & Noble, K. (2021). *Feasibility of Assessing Brain Activity using Mobile, In-home Collection of Electroencephalography: Methods and Analysis*. PsyArXiv. https://doi.org/10.31234/osf.io/rckfa

Troller-Renfree, S. V., Brito, N. H., Desai, P. M., Leon-Santos, A. G., Wiltshire, C. A., Motton, S. N., Meyer, J. S., Isler, J., Fifer, W. P., & Noble, K. G. (2020). Infants of mothers with higher physiological stress show alterations in brain function. *Developmental Science*, *23*(6), e12976. https://doi.org/10.1111/desc.12976

Troller-Renfree, S. V., Costanzo, M. A., Duncan, G. J., Magnuson, K., Gennetian, L. A., Yoshikawa, H., Halpern-Meekin, S., Fox, N. A., & Noble, K. G. (2022). The impact of a poverty reduction intervention on infant brain activity. *Proceedings of the National Academy of Sciences*, *119*(5), e2115649119. https://doi.org/10.1073/pnas.2115649119

Troller-Renfree, S. V., Morales, S., Leach, S. C., Bowers, M. E., Debnath, R., Fifer, W. P., Fox, N. A., & Noble, K. G. (2021). Feasibility of assessing brain activity using mobile, in-home collection of electroencephalography: Methods and analysis. *Developmental Psychobiology*, *63*(6), e22128. https://doi.org/10.1002/dev.22128

Turkheimer, E., Haley, A., Waldron, M., D'Onofrio, B., & Gottesman, I. I. (2003). Socioeconomic Status Modifies Heritability of IQ in Young Children. *Psychological Science*, *14*(6), 623–628. https://doi.org/10.1046/j.0956-7976.2003.psci_1475.x

Ursache, A., & Noble, K. G. (2016). Neurocognitive development in socioeconomic context: Multiple mechanisms and implications for measuring socioeconomic status. *Psychophysiology*, *53*(1), 71–82. https://doi.org/10.1111/psyp.12547

Ursache, A., Noble, K. G., & the Pediatric Imaging, Neurocognition and Genetics Study. (2016). Socioeconomic status, white matter, and executive function in children. *Brain and Behavior*, *6*(10), e00531. https://doi.org/10.1002/brb3.531

van Baar, A. L., de Jong, M., Maat, M., Hooge, I. T. C., Bogičević, L., & Verhoeven, M. (2020). Reliability and Validity of the Utrecht Tasks for Attention in Toddlers Using Eye Tracking (UTATE). *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.01179

van der Velde, B., Haartsen, R., & Kemner, C. (2019). Test-retest reliability of EEG network characteristics in infants. *Brain and Behavior*, *9*(5), e01269. https://doi.org/10.1002/brb3.1269

van der Velde, B., White, T., & Kemner, C. (2021). The emergence of a theta social brain network during infancy. *NeuroImage*, *240*, 118298. https://doi.org/10.1016/j.neuroimage.2021.118298

van Liempd, H. (Ine) M. J. A., Oudgenoeg-Paz, O., Fukkink, R. G., & Leseman, P. P. M. (2018). Young children's exploration of the indoor playroom space in center-based childcare. *Early Childhood Research Quarterly*, *43*, 33–41. https://doi.org/10.1016/j.ecresq.2017.11.005

van Noordt, S., Heffer, T., & Willoughby, T. (2022). A developmental examination of medial frontal theta dynamics and inhibitory control. *NeuroImage*, *246*, 118765. https://doi.org/10.1016/j.neuroimage.2021.118765

Van Schijndel, T. J. P., Franse, R. K., & Raijmakers, M. E. J. (2010). The Exploratory Behavior Scale: Assessing young visitors' hands-on behavior in science museums. *Science Education*, *94*(5), 794–809. https://doi.org/10.1002/sce.20394

Virani, A., Duffett-Leger, L., & Letourneau, N. (2021). Parents' use of mobile applications in the first year of parenthood: A narrative review of the literature. *Health Technology*, *5*, 14–14. https://doi.org/10.21037/ht-20-28

von Stumm, S., & Plomin, R. (2015). Socioeconomic status and the growth of intelligence from infancy through adolescence. *Intelligence*, *48*, 30–36. https://doi.org/10.1016/j.intell.2014.10.002

Vrantsidis, D. M., Clark, C. A. C., Chevalier, N., Espy, K. A., & Wiebe, S. A. (2020). Socioeconomic status and executive function in early childhood: Exploring proximal mechanisms. *Developmental Science*, *23*(3), e12917. https://doi.org/10.1111/desc.12917

Walpole, M. (2003). Socioeconomic Status and College: How SES Affects College Experiences and Outcomes. *Fall*, *27*, 45–73. https://doi.org/10.1353/rhe.2003.0044

Wass, S. V., Noreika, V., Georgieva, S., Clackson, K., Brightman, L., Nutbrown, R., Covarrubias, L. S., & 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

Wass, S. V., Smith, C. G., Daubney, K. R., Suata, Z. M., Clackson, K., Begum, A., & Mirza, F. U. (2019). Influences of environmental stressors on autonomic function in 12-month-old infants: Understanding early common pathways to atypical emotion regulation and cognitive performance. *Journal of Child Psychology and Psychiatry*, *60*(12), 1323–1333. https://doi.org/10.1111/jcpp.13084

Webb, S. J., Shic, F., Murias, M., Sugar, C. A., Naples, A. J., Barney, E., Borland, H., Hellemann, G., Johnson, S., Kim, M., Levin, A. R., Sabatos-DeVito, M., Santhosh, M., Senturk, D., Dziura, J., Bernier, R. A., Chawarska, K., Dawson, G., Faja, S., … Yuan, A. (2020). Biomarker Acquisition and Quality Control for Multi-Site Studies: The Autism Biomarkers Consortium for Clinical Trials. *Frontiers in Integrative Neuroscience*, *13*. https://doi.org/10.3389/fnint.2019.00071

Weisler, A., & McCall, R. R. (1976). Exploration and play: Resume and redirection. *American Psychologist*, *31*(7), 492–508. https://doi.org/10.1037/0003-066X.31.7.492

Wellman, H. M., Lopez-Duran, S., LaBounty, J., & Hamilton, B. (2008). Infant attention to intentional action predicts preschool theory of mind. *Developmental Psychology*, *44*(2), 618–623. https://doi.org/10.1037/0012-1649.44.2.618

Westphal, K. P., Grözinger, B., Diekmann, V., & Kornhuber, H. H. (1993). EEG-blocking before and during voluntary movements: Differences between the eyes-closed and the eyes-open condition. *Archives Italiennes De Biologie*, *131*(1), 25–35.

Wiebe, S. A., Espy, K. A., & Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Developmental Psychology*, *44*(2), 575–587. https://doi.org/10.1037/0012-1649.44.2.575

Wild, K. T., Betancourt, L. M., Brodsky, N. L., & Hurt, H. (2013). The effect of socioeconomic status on the language outcome of preterm infants at toddler age. *Early Human Development*, *89*(9), 743–746. https://doi.org/10.1016/j.earlhumdev.2013.05.008

Wolke, D., Jaekel, J., Hall, J., & Baumann, N. (2013). Effects of Sensitive Parenting on the Academic Resilience of Very Preterm and Very Low Birth Weight Adolescents. *Journal of Adolescent Health*, *53*(5), 642–647. https://doi.org/10.1016/j.jadohealth.2013.06.014

World Health Organization. (2022). *Optimizing brain health across the life course: WHO position paper*. https://www.who.int/publications-detail-redirect/9789240054561

Xia, R., Zhao, X., Liu, Y., Dou, Y., Shu, Z., Ding, X., Zhou, X., Han, J., & Zhao, X. (2024). Attention network training promotes selective attention of children with low socioeconomic status. *Journal of Experimental Child Psychology*, *239*, 105807. https://doi.org/10.1016/j.jecp.2023.105807

Xie, W., Mallin, B. M., & Richards, J. E. (2018). Development of infant sustained attention and its relation to EEG oscillations: An EEG and cortical source analysis study. *Developmental Science*, *21*(3), e12562. https://doi.org/10.1111/desc.12562

Xu, K., Torgrimson, S. J., Torres, R., Lenartowicz, A., & Grammer, J. K. (2022). EEG Data Quality in Real-World Settings: Examining Neural Correlates of Attention in School-Aged Children. *Mind, Brain, and Education*, *16*(3), 221–227. https://doi.org/10.1111/mbe.12314

Xu, R. Y., Chow, C. H. T., Poole, K. L., Sriranjan, J., Mirabelli, J., Moffat, G., Van Lieshout, R. J., Buckley, N., & Schmidt, L. A. (2023). Mobile Quantitative EEG Assessment and Children's Preoperative Anxiety: A Psychometric and Feasibility Study. *Journal of Technology in Behavioral Science*, *8*(1), 47–56. https://doi.org/10.1007/s41347-022-00285-z

Zawati, M. H., Parry, D., & Knoppers, B. (2014). *The Best Interests of the Child and the Return of Results in Genetic Research: International Comparative Perspectives* (SSRN Scholarly Paper No. 2632721). https://papers.ssrn.com/abstract=2632721

Appendices

APPENDIX A. CUT-OFF ANALYSES FOR CHAPTER 2 CONDITION ANALYSES AFTER CUT-OFF CRITERIA WERE APPLIED (SEE SECTION 2.3.1.3)

For each task, participants were excluded according to cut-off criteria (Table A 1); condition effect analyses as reported in the main text were then repeated. All statistical tests were conducted as in the main text; results are presented in tables here for succinctness. All tasks showed the same pattern of results before and after exclusion criteria were applied.

Table A 1

Exclusion criteria which were applied to each task prior to cut-off analyses being performed

Task	Cut-off criteria
Gap	< 5 valid trials per condition
Non-social Contingency	< 5 valid trials per condition
Reversal Learning	< 2 valid trials per phase
Working Memory	< 10 valid trials
Visual Search	< 3 valid trials per condition
Face pop-out	< 3 valid trials
Dancing Ladies	< 3 valid trials
Fifty Faces	< 20% trials valid

Appendix A.1 Gap overlap task

Table A 2

Effects from repeated measures ANOVA comparing reaction times in gap, overlap and baseline conditions

Effect	*F*	LB	UB	*p*	η_p^2
Overall condition	1191.41	1.69	508.89	< 0.001	0.80
Facilitation	1225.06	1	301	< 0.001	0.80
Disengagement	436.81	1	301	< 0.001	0.59

Table A 3

Mean, standard deviation and n for gap, overlap and baseline conditions

Condition	*M*	*SD*	*n*
Gap	554.25	11.05	302
Baseline	574.60	12.36	302
Overlap	590.11	17.43	302

Appendix A.2 Non-social Contingency

Participant number did not change, analyses the same as in main text.

Appendix A.3 Reversal Learning Task

Table A 4

Mean, standard deviation and n for key variables in the pre-switch phase

	Mean	*SD*	*n*
CC-Pre-Acc	0.71	0.26	321
CC-Pre-SRT	689.87	204.32	321

Of the group of 321 participants who provided at least 2 valid trials of the learning phase of the cognitive control task, 203 participants proceeded to the reversal condition. Of these, 194 participants (60.4% of the learning sample) had at least two valid trails in the reversal condition and are included in the current analyses. In total 312 participants provided at least two valid trials in each of the phases they took part in.

During the pre-switch phase of the cognitive control task, the mean proportion of trials in which participants correctly anticipated animation (CC-Pre-Acc) was approximately 71%; *M* = 0.71, *SD* = 0.26, *N* = 312. The mean reaction time for participants to select an AOI (CC-Pre-SRT) was 691.82, *SD* = 205.87, *N* = 312.

Table A 5

Results from repeated-measures t-test comparisons between pre-switch and post-switch phase for proportion of trials in which participants correctly anticipated animation and reaction time

	t	*df*	*p*	*d*	LB	UB
CC-Pre-Acc vs CC-Post-Acc	2.72	193	0.007	0.29	0.02	0.11
CC-Pre-SRT vs CC-Post-SRT	-0.76	193	0.45	-0.05	-0.04	0.02

Table A 6

Mean, standard deviation and n for proportion of trials in which participants correctly anticipated animation and reaction time in pre-switch and post-switch phases

	M	*SD*	*n*
CC-Pre-Acc	0.84	0.16	194
CC-Post-Acc	0.78	0.25	194
CC-Pre-SRT	665.15	185.65	194
CC-Post-SRT	665.86	180.35	194

Table A 7

Results from independent t-tests comparisons time between groups who did and didn't do the post-switch phase for proportion of trials in which participants correctly anticipated animation and reaction

	t	*df*	*p*	*d*	LB	UB
CC-Pre-Acc	-13.36	175.54	< 0.001	1.62	-.40	-.30
CC-Pre-SRT	3.96	213.13	< 0.001	0.47	.05	.15

Table A 8

Mean, standard deviation and n for proportion of trials in which participants correctly anticipated animation and reaction time for groups who did and didn't do the post-switch phase

Variable	Group	*M*	*SD*	*n*
Proportion of trials correctly anticipated	Did reversal condition	0.84	0.16	194
	Didn't do reversal condition	0.49	0.26	118
Reaction time	Did reversal condition	655.15	185.65	194
	Didn't do reversal condition	752.09	223.40	118

Appendix A.4 Working Memory

Table A 9

Mean, standard deviation and n for reaction times in all trials, in correct and in incorrect trials, and the proportion of trials in which participants correctly anticipated animation

	M	*SD*	*n*
Mean RT on all trials	696.43	169.17	311
Mean RT in correct trials	714.90	200.67	304
Mean RT in incorrect trials	707.49	195.62	304
WM-Acc	0.48	0.19	311

Table A 10

Results from a one-sample t-test comparing the proportion of trials in which participants correctly anticipated animation to chance and from a paired samples t-test between reaction time in correct and incorrect trials

	t	*df*	*p*	*d*	LB	UB
WM-Acc vs. chance	-1.53	310	0.13	-0.11	-.04	0.01
WM-SRT-Acc vs WM-SRT-Inacc	0.61	303	0.54	0.04	-16.5	31.3

Appendix A.5 Visual Search Task

Table A 11

Results from repeated-measures ANOVA comparing reaction time across three conditions of the visual search task

Effect		*F*	*df (error)*	*df (condition)*	*p*	η_p^2
Condition		157.60	1.76	447.03	< 0.001	0.38
VS-S9-SRT	VS-C9-SRT	251.77	1	254	< 0.001	0.50
VS-C13-SRT	VS-C9-SRT	5.59	1	254	0.02	0.02

Table A 12

Mean, standard deviation and n for reaction times across three conditions of the visual search task

Condition	*M*	*SD*	*n*
VS-S9-SRT	940.20	241.53	255
VS-C9-SRT	1317.90	369.51	255
VS-C13-SRT	1393.40	372.73	255

Table A 13

Results from repeated-measures ANOVA comparing proportion of correct trials across three conditions of the visual search task

Effect		*F*	*df (error)*	*df (condition)*	*p*	η_p^2
Condition		320.02	2	634	< 0.001	0.50
VS-S9-Acc	VS-C9-Acc	411.31	1	317	< 0.001	0.57
VS-C13-Acc	VS-C9-Acc	21.74	1	317	< 0.001	0.06

Table A 14

Mean, standard deviation and n for proportion of correct trials across three conditions of the visual search task

Condition	*M*	*SD*	*n*
VS-S9-Acc	0.76	0.24	318
VS-C9-Acc	0.46	0.26	318
VS-C13-Acc	0.39	0.26	318

Appendix A.6 Pop-out task

Table A 15

Results from repeated-measures ANOVA comparing proportion looking to face, car and noise AOIs

Effect	*F*	*df (error)*	*df (condition)*	*p*	η_p^2
Condition	172.45	1.38	338.94	< 0.001	0.41
Face vs. car	65.36	1	245	< 0.001	0.21
Face vs noise	637.34	1	245	< 0.001	0.72

Table A 16

Mean, standard deviation and n for proportion looking to face, car and noise AOIs

	M	*SD*	*n*
Face	0.32	0.14	246
Car	0.20	0.14	246
Noise	0.10	0.05	246

Table A 17

Results from repeated-measures ANOVA comparing peak look duration to face, car, and noise AOIs

Effect	*F*	*df (error)*	*df (condition)*	*p*	η_p^2
Condition	95.36	1.42	347.18	< 0.001	0.28
Face vs. car	8.42	1	245	0.004	0.03
Face vs noise	374.94	1	245	< 0.001	0.61

Table A 18

Mean, standard deviation and n for peak look duration to face, car, and noise AOIs

	M	*SD*	*n*
Face	1.14	0.48	246
Car	0.98	0.60	246
Noise	0.56	0.19	246

Table A 19

Results from one-sample t-tests comparing proportion of looking to faces and proportion of first looks to faces to chance level

	t	*df*	*p*	*d*	LB	UB
Pop-Face-Pct vs chance	14.26	245	< 0.001	0.92	0.11	0.14
Pop-Face-First vs chance	26.83	245	< 0.001	1.70	0.36	0.42

Table A 20

Mean, standard deviation and n for proportion of looking to faces and proportion of first looks to faces

	M	*SD*	*n*
Pop-Face-Pct	0.32	0.13	246
Pop-Face-First	0.59	0.23	246

Appendix A.7 Dancing ladies

A 2 x 2 repeated-measures ANOVA showed that peak look to faces (*M* = 0.62, *SE* = 0.02) was significantly higher than to objects (*M* = 0.51, *SE* = 0.01), (F(1, 222) = 24.76, p < 0.001, η_p^2 = 0.10), and peak look to faces was significantly greater in social (*M* = 0.75, *SD* = 0.40) than scrambled (*M* = 0.50, *SD* = 0.27), (t(222) = 11.12, p < 0.001, d = 0.73).

Proportion looking to faces (*M* = 0.14, se = 0.01) was also significantly higher than to objects (*M* = 0.10, se = 0.002), (F(1, 225) = 33.08, p < 0.001, η_p^2 = 0.13), and proportion looking to faces was significantly higher in social (*M* = 0.20, SD = 0.13) than scrambled conditions (*M* = 0.07, SD = 0.06), (t(225) = 17.31, p < 0.001, d = 1.28)

Appendix A.8 Fifty faces

Table A 21

Results from paired t-tests between peak look duration and proportion of looking to faces compared to background people

	t	*df*	*p*	*d*	LB	UB
Peak look duration	26.21	222	< 0.001	2.48	2.99	3.47
Proportion of looking	62.14	234	< 0.001	6.02	0.54	0.58

Table A 22

Mean, standard deviation and n for peak look duration and proportion of looking to faces and background people

	Condition	*M*	*SD*	*n*
Peak look duration	Faces	3.69	1.81	223
	Background people	0.46	0.34	223
Proportion of looking	Faces	0.59	0.13	235
	Background people	0.03	0.02	235

APPENDIX B. ADDITIONAL ANALYSES FOR CHAPTER 3

To assess the robustness of findings, particularly in light of findings that 20 segments were needed for good reliability and 40 for excellent reliability (Troller-Renfree et al., 2021), analyses were repeated using a minimum of 30 usable EEG segments per condition. Using a stricter minimum number of trials resulted in broadly the same pattern of findings in condition comparisons, and did not significantly impact test-retest reliabilities. Results from these analyses are reported here.

Appendix B.1 Condition differences

Appendix B.1a Absolute

Table B 1

*F-value, p-value and partial eta squared effect size for ANOVAs on average relative power over the whole video for test and retest sessions. Significant effects are indicated with a **

	TEST			RETEST		
	F	*p*	η^2 (partial)	*F*	*p*	η^2 (partial)
Signal	742.25	<0.001*	0.46	1007.82	<0.001*	0.54
Region	44.77	<0.001*	0.03	83.31	<0.001*	0.04
Video number	0.35	0.708	0.00	0.23	0.791	0.00
Video condition	4.35	0.037*	0.00	0.95	0.331	0.00
Signal x Region	6.33	0.012*	0.00	10.61	0.001*	0.01
Signal x Video number	2.84	0.059	0.00	0.91	0.403	0.00
Region x Video number	0.26	0.771	0.00	0.09	0.916	0.00
Signal x Video condition	0.01	0.911	0.00	0.68	0.411	0.00

Region x Video condition	0.48	0.487	0.00	1.09	0.297	0.00
Video number x Video condition	0.04	0.959	0.00	0.15	0.864	0.00
Signal x Region x Video number	0.03	0.966	0.00	0.02	0.985	0.00
Signal x Region x Video condition	0.01	0.922	0.00	0.00	0.979	0.00
Signal x Video condition x Video number	0.06	0.943	0.00	0.10	0.901	0.00
Region x Video condition x Video number	0.11	0.900	0.00	0.01	0.987	0.00
Signal x Region x Video number x Video condition	0.05	0.951	0.00	0.11	0.898	0.00

Table B 2

Mean, standard deviation and N for average relative power over the whole video for test and retest sessions

Signal	*Region*	*Video number*	*Video condition*	TEST			RETEST		
				M	*SD*	*N*	*M*	*SD*	*N*
Theta	Frontal	First	Non-social	1.74	0.39	34	1.69	0.40	34
			Social	1.76	0.39	33	1.71	0.42	33
		Second	Non-social	1.65	0.31	33	1.67	0.39	33
			Social	1.71	0.38	36	1.63	0.37	36
		Third	Non-social	1.65	0.38	31	1.66	0.43	31
			Social	1.70	0.40	32	1.60	0.44	32
	Posterior	First	Non-social	1.91	0.42	34	1.98	0.30	34
			Social	2.00	0.35	33	2.03	0.35	33
		Second	Non-social	1.90	0.40	33	1.97	0.30	33
			Social	1.97	0.51	36	1.97	0.34	36
		Third	Non-social	1.91	0.38	31	1.93	0.34	31
			Social	1.97	0.46	32	1.97	0.34	32
Alpha	Frontal	First	Non-social	2.42	0.29	34	2.54	0.32	34
			Social	2.47	0.27	33	2.55	0.37	33
		Second	Non-social	2.48	0.34	33	2.54	0.34	33
			Social	2.52	0.34	36	2.57	0.38	36
		Third	Non-social	2.53	0.34	31	2.54	0.34	31

		Social	2.53	0.30	32	2.56	0.35	32
Posterior	First	Non-social	2.48	0.40	34	2.63	0.32	34
		Social	2.58	0.40	33	2.73	0.35	33
	Second	Non-social	2.60	0.35	33	2.67	0.36	33
		Social	2.65	0.43	36	2.74	0.36	36
	Third	Non-social	2.61	0.36	31	2.69	0.33	31
		Social	2.68	0.43	32	2.74	0.35	32

Table B 3

Mean, standard deviation and N for average absolute alpha and theta power for test and retest sessions

Signal	TEST					RETEST				
	M	*SE*	*N*	*LCL*	*UCL*	*M*	*SE*	*N*	*LCL*	*UCL*
Alpha	1.82	0.02	412	1.79	1.86	1.82	0.02	398	1.78	1.85
Theta	2.55	0.02	412	2.51	2.58	2.62	0.02	398	2.59	2.66

Table B 4

Mean, standard deviation and N for average absolute frontal and posterior power for test and retest sessions

Region	TEST					RETEST				
	M	*SE*	*N*	*LCL*	*UCL*	*M*	*SE*	*N*	*LCL*	*UCL*
Frontal	2.10	0.02	412	2.06	2.13	2.10	0.02	398	2.07	2.14
Posterior	2.27	0.02	412	2.24	2.31	2.34	0.02	398	2.30	2.37

Table B 5

Mean, standard deviation and N for average absolute alpha and theta power in frontal and posterior regions for test and retest sessions

		TEST					RETEST				
Signal	*Region*	*M*	*SD*	*N*	*LCL*	*UCL*	*M*	*SD*	*N*	*LCL*	*UCL*
Alpha	Frontal	1.70	0.03	800	1.63	1.78	1.66	0.03	772	1.59	1.73
	Posterior	1.95	0.03	800	1.87	2.02	1.97	0.03	772	1.90	2.05
Theta	Frontal	2.49	0.03	800	2.42	2.57	2.55	0.03	772	2.48	2.62
	Posterior	2.60	0.03	800	2.53	2.68	2.70	0.03	772	2.63	2.77

Table B 6

*t-value, standard error, degrees of freedom, and p-value for pairwise comparisons between absolute alpha and theta power in frontal and posterior regions in the test and retest sessions. Significant effects are indicated with a **

contrast	TEST				RETEST			
	t-value	SE	df	*p-value*	*t*-value	SE	df	*p-value*
frontal alpha - posterior alpha	-6.46	0.04	800	<0.001*	-8.73	0.04	772	<0.001*
frontal alpha - frontal theta	-20.96	0.04	800	<0.001*	-24.71	0.04	772	<0.001*
frontal alpha - posterior theta	-23.87	0.04	800	<0.001*	-28.87	0.04	772	<0.001*
posterior alpha - frontal theta	-14.5	0.04	800	<0.001*	-15.98	0.04	772	<0.001*

posterior alpha - posterior theta	-17.41	0.04	800	<0.001*	-20.13	0.04	772	<0.001*
frontal theta - posterior theta	-2.91	0.04	800	0.019*	-4.15	0.04	772	<0.001*

Appendix B.1b Relative

Table B 7

*F-value, p-value and partial eta squared effect size for ANOVAs on average relative power over the whole video for test and retest sessions. Significant effects are indicated by **

	TEST			RETEST		
	F	*p*	η^2 (partial)	*F*	*p*	η^2 (partial)
Signal	677.70	<.001*	0.44	1046.61	<.001*	0.56
Region	21.16	<.001*	0.01	11.60	0.001*	0.01
Video number	0.93	0.394	0.00	2.15	0.117	0.00
Video condition	6.71	0.010*	0.00	4.60	0.032*	0.00
Signal x Region	14.84	<.001*	0.01	13.44	<.001*	0.01
Signal x Video number	5.51	0.004	0.01	1.58	0.206	0.00
Region x Video number	0.11	0.898	0.00	0.29	0.751	0.00
Signal x Video condition	0.00	0.986	0.00	1.03	0.310	0.00
Region x Video condition	1.46	0.227	0.00	0.00	0.970	0.00

Video number x Video condition	0.09	0.910	0.00	0.19	0.826	0.00
Signal x Region x Video number	0.07	0.936	0.00	0.13	0.877	0.00
Signal x Region x Video condition	0.03	0.871	0.00	0.67	0.413	0.00
Signal x Video condition x Video number	0.22	0.806	0.00	0.06	0.942	0.00
Region x Video condition x Video number	0.29	0.749	0.00	0.19	0.828	0.00
Signal x Region x Video number x Video condition	0.12	0.888	0.00	0.57	0.564	0.00

Table B 8

Mean, standard deviation and N for average relative power over the whole video for test and retest sessions

Signal	*Region*	*Video number*	*Video condition*	TEST			RETEST		
				M	*SD*	*N*	*M*	*SD*	*N*
Alpha	Frontal	First	Non-social	0.30	0.08	32	0.28	0.07	34
			Social	0.31	0.08	34	0.29	0.07	33
		Second	Non-social	0.28	0.07	32	0.27	0.07	33
			Social	0.29	0.07	38	0.27	0.07	36
		Third	Non-social	0.28	0.07	33	0.27	0.08	31

			Social	0.28	0.08	37	0.26	0.07	32
	Posterior	First	Non-social	0.29	0.06	32	0.28	0.06	34
			Social	0.31	0.06	34	0.29	0.06	33
		Second	Non-social	0.27	0.06	32	0.27	0.06	33
			Social	0.29	0.07	38	0.28	0.07	36
		Third	Non-social	0.27	0.06	33	0.26	0.06	31
			Social	0.28	0.07	37	0.27	0.06	32
Theta	Frontal	First	Non-social	0.41	0.07	32	0.44	0.06	34
			Social	0.43	0.06	34	0.45	0.07	33
		Second	Non-social	0.42	0.07	32	0.42	0.07	33
			Social	0.43	0.07	38	0.45	0.08	36
		Third	Non-social	0.43	0.06	33	0.42	0.05	31
			Social	0.43	0.06	37	0.45	0.05	32
	Posterior	First	Non-social	0.37	0.06	32	0.39	0.05	34
			Social	0.39	0.05	34	0.41	0.06	33
		Second	Non-social	0.38	0.06	32	0.40	0.06	33
			Social	0.40	0.06	38	0.41	0.07	36
		Third	Non-social	0.38	0.05	33	0.41	0.05	31
			Social	0.40	0.06	37	0.41	0.06	32

Table B 9

Mean, standard deviation and N for average relative alpha and theta power for test and retest sessions

Signal	TEST					RETEST				
	M	*SE*	*N*	*LCL*	*UCL*	*M*	*SE*	*N*	*LCL*	*UCL*
Alpha	0.29	0.00	412	0.28	0.29	0.27	0.00	398	0.27	0.28
Theta	0.41	0.00	412	0.40	0.41	0.42	0.00	398	0.42	0.43

Table B 10

Mean, standard deviation and N for average relative frontal and posterior power for test and retest sessions

Region	TEST					RETEST				
	M	*SE*	*N*	*LCL*	*UCL*	*M*	*SE*	*N*	*LCL*	*UCL*
Frontal	0.36	0.00	412	0.35	0.36	0.36	0.00	398	0.35	0.36
Posterior	0.34	0.00	412	0.33	0.34	0.34	0.00	398	0.33	0.35

Table B 11

Mean, standard deviation and N for average relative alpha and theta power in frontal and posterior regions for test and retest session

Signal	*Region*	TEST					RETEST				
		M	*SD*	*N*	*LCL*	*UCL*	*M*	*SD*	*N*	*LCL*	*UCL*
Alpha	Frontal	0.29	0.00	800	0.27	0.30	0.27	0.00	772	0.26	0.29
	Posterior	0.28	0.00	800	0.27	0.30	0.27	0.00	772	0.26	0.29
Theta	Frontal	0.43	0.00	800	0.41	0.44	0.44	0.00	772	0.42	0.45
	Posterior	0.39	0.00	800	0.37	0.40	0.41	0.00	772	0.39	0.42

Table B 12

*t-value, standard error, degrees of freedom, and p-value for pairwise comparisons between relative alpha and theta power in frontal and posterior regions in the test and retest sessions. Significant effects are indicated with a **

contrast	TEST				RETEST			
	t-value	SE	df	*p-value*	*t*-value	SE	df	*p-value*
frontal alpha - posterior alpha	0.57	0.01	800	0.940	-0.17	0.01	772	0.998
frontal alpha - frontal theta	-21.03	0.01	800	<0.001*	-25.41	0.01	772	<0.001*
frontal alpha - posterior theta	-15.02	0.01	800	<0.001*	-20.46	0.01	772	<0.001*
posterior alpha - frontal theta	-21.61	0.01	800	<0.001*	-25.25	0.01	772	<0.001*
posterior alpha - posterior theta	-15.59	0.01	800	<0.001*	-20.29	0.01	772	<0.001*
frontal theta - posterior theta	6.02	0.01	800	<0.001*	4.95	0.01	772	<0.001*

Appendix B.2 Halves difference

Appendix B.2a Absolute

Table B 13

*F-value, p-value and partial eta squared effect size for ANOVAs comparing condition effects for average absolute power over the first and second half of video viewing for test and retest sessions. Significant effects are indicated by **

TEST	RETEST

	F	*p*	η² (partial)	*F*	*p*	η² (partial)
Signal	1092.46	<.001*	0.46	1967.05	<.001*	0.51
Region	61.98	<.001*	0.03	185.22	<.001*	0.05
Video number	1.04	0.352	0.00	0.62	0.539	0.00
Video condition	3.20	0.074	0.00	4.00	0.046*	0.00
Half	5.18	0.023*	0.00	1.92	0.166	0.00
Signal x Region	10.59	0.001*	0.00	19.16	<.001*	0.00
Signal x Video number	5.26	0.005*	0.00	1.94	0.144	0.00
Region x Video number	0.57	0.565	0.00	0.12	0.884	0.00
Signal x Video condition	0.21	0.647	0.00	0.88	0.348	0.00
Region x Video condition	0.25	0.616	0.00	1.97	0.160	0.00
Video number x Video condition	1.86	0.155	0.00	0.37	0.688	0.00
Signal x Half	1.96	0.162	0.00	0.27	0.600	0.00
Region x Half	0.16	0.686	0.00	0.28	0.600	0.00
Video number x Half	0.12	0.889	0.00	0.06	0.942	0.00
Video condition x Half	0.20	0.656	0.00	0.12	0.728	0.00
Signal x Region x Video number	0.06	0.942	0.00	0.02	0.980	0.00
Signal x Region x Video condition	0.02	0.900	0.00	0.00	0.983	0.00
Signal x Video number x Video condition	0.56	0.571	0.00	0.08	0.920	0.00

Region x Video number x Video condition	0.02	0.976	0.00	0.22	0.804	0.00
Signal x Region x Video Half	0.04	0.834	0.00	0.08	0.778	0.00
Signal x Video number x Half	0.05	0.949	0.00	0.02	0.979	0.00
Region x Video number x Half	0.09	0.912	0.00	0.06	0.944	0.00
Signal x Video condition x Half	0.11	0.740	0.00	0.62	0.431	0.00
Region x Video condition x Half	0.42	0.516	0.00	0.07	0.795	0.00
Video number x Video condition x Half	0.08	0.927	0.00	0.11	0.895	0.00
Signal x Region x Video number x Video condition	0.03	0.966	0.00	0.24	0.785	0.00
Signal x Region x Video number x Half	0.06	0.940	0.00	0.07	0.928	0.00
Signal x Region x Video condition x Half	0.01	0.940	0.00	0.05	0.830	0.00
Signal x Video number x Video condition x Half	0.02	0.976	0.00	0.29	0.751	0.00
Region x Video number x Video condition x Half	0.42	0.660	0.00	0.13	0.881	0.00
Signal x Region x Video number x Video condition x Half	0.04	0.956	0.00	0.04	0.958	0.00

Table B 14

Mean, standard deviation and N for average absolute power over video halves for test and retest sessions

Signal	*Region*	*Video number*	*Video condition*	*Video half*	TEST			RETEST		
					M	*SD*	*N*	*M*	*SD*	*N*
Alpha	Frontal	First	Non-social	First	1.73	0.45	29	1.65	0.41	37
				Second	1.76	0.40	29	1.67	0.42	37
			Social	First	1.68	0.37	27	1.67	0.44	37
				Second	1.72	0.42	27	1.69	0.50	37
		Second	Non-social	First	1.64	0.37	25	1.65	0.43	36
				Second	1.62	0.33	25	1.63	0.41	36
			Social	First	1.70	0.43	24	1.63	0.37	37
				Second	1.79	0.40	24	1.63	0.41	37
		Third	Non-social	First	1.60	0.39	25	1.62	0.46	34
				Second	1.64	0.34	25	1.60	0.47	34
			Social	First	1.67	0.40	26	1.58	0.50	34
				Second	1.69	0.36	26	1.58	0.45	34
	Posterior	First	Non-social	First	1.96	0.36	29	1.95	0.31	37
				Second	1.95	0.40	29	1.96	0.31	37
			Social	First	1.94	0.34	27	1.99	0.36	37
				Second	1.94	0.37	27	2.06	0.34	37
		Second	Non-social	First	1.81	0.45	25	1.95	0.32	36

				Second	1.88	0.42	25	1.98	0.32	36
			Social	First	1.98	0.55	24	1.96	0.36	37
				Second	1.96	0.51	24	1.99	0.34	37
		Third	Non-social	First	1.90	0.32	25	1.89	0.35	34
				Second	1.90	0.32	25	1.92	0.35	34
			Social	First	1.97	0.35	26	1.95	0.38	34
				Second	1.97	0.36	26	1.97	0.33	34
Theta	Frontal	First	Non-social	First	2.38	0.32	29	2.48	0.36	37
				Second	2.45	0.31	29	2.57	0.35	37
			Social	First	2.34	0.26	27	2.55	0.45	33
				Second	2.47	0.30	27	2.52	0.38	37
		Second	Non-social	First	2.50	0.38	25	2.48	0.33	36
				Second	2.52	0.39	25	2.55	0.42	36
			Social	First	2.47	0.39	24	2.56	0.41	37
				Second	2.60	0.44	24	2.58	0.39	37
		Third	Social	First	2.45	0.35	25	2.51	0.35	34
				Second	2.52	0.32	25	2.54	0.38	34
			Non-social	First	2.46	0.30	26	2.54	0.45	34
				Second	2.56	0.30	26	2.55	0.39	34
	Posterior	First	Non-social	First	2.47	0.42	29	2.58	0.34	37
				Second	2.53	0.40	29	2.65	0.36	37
			Social	First	2.44	0.37	27	2.74	0.37	37

			Second	2.57	0.44	27	2.76	0.35	37
	Second	Non-social	First	2.53	0.40	25	2.65	0.32	36
			Second	2.64	0.39	25	2.71	0.42	36
		Social	First	2.59	0.43	24	2.76	0.37	37
			Second	2.66	0.43	24	2.74	0.38	37
	Third	Non-social	First	2.56	0.37	25	2.68	0.35	34
			Second	2.60	0.33	25	2.71	0.34	34
		Social	First	2.65	0.37	26	2.70	0.36	34
			Second	2.67	0.40	26	2.78	0.33	34

Table B 15

Mean, standard deviation and N for average absolute power in the first and second half of videos in the test session

Half	TEST				
	M	*SE*	*Df*	*LCL*	*UCL*
First	2.14	0.02	1200	2.11	2.18
Second	2.19	0.02	1200	2.16	2.23

Appendix B.2b Relative

Table B 16

*F-value, p-value and partial eta squared effect size for ANOVAs comparing condition effects for average relative power over the first and second half of video viewing for test and retest sessions. Significant effects are indicated with a **

	TEST			RETEST		
	F	*p*	η² (partial)	*F*	*p*	η² (partial)
Signal	890.21	<.001*	0.41	2041.56	<.001*	0.53
Region	26.11	<.001*	0.01	31.50	<.001*	0.01
Video number	0.18	0.838	0.00	7.17	0.001*	0.00
Video condition	9.67	0.002*	0.00	17.43	<.001*	0.00
Half	1.43	0.233	0.00	1.34	0.247	0.00
Signal x Region	20.16	<.001*	0.01	33.11	<.001*	0.01
Signal x Video number	9.61	<.001*	0.01	1.91	0.149	0.00
Region x Video number	0.04	0.961	0.00	0.27	0.760	0.00
Signal x Video condition	0.51	0.477	0.00	3.21	0.073	0.00
Region x Video condition	1.12	0.290	0.00	0.92	0.337	0.00
Video number x Video condition	1.46	0.232	0.00	0.08	0.920	0.00
Signal x Half	4.13	0.042*	0.00	0.69	0.406	0.00
Region x Half	0.21	0.647	0.00	0.75	0.386	0.00
Video number x Half	0.16	0.852	0.00	0.09	0.915	0.00
Video condition x Half	0.74	0.391	0.00	0.02	0.886	0.00
Signal x Region x Video number	0.08	0.925	0.00	0.02	0.982	0.00
Signal x Region x Video condition	0.01	0.929	0.00	0.09	0.768	0.00
Signal x Video number x Video condition	1.67	0.189	0.00	0.19	0.825	0.00

Region x Video number x Video condition	0.34	0.710	0.00	0.06	0.946	0.00
Signal x Region x Video Half	0.03	0.870	0.00	0.04	0.836	0.00
Signal x Video number x Half	0.17	0.844	0.00	0.11	0.896	0.00
Region x Video number x Half	0.27	0.765	0.00	0.12	0.884	0.00
Signal x Video condition x Half	0.22	0.642	0.00	0.42	0.517	0.00
Region x Video condition x Half	0.01	0.925	0.00	0.08	0.779	0.00
Video number x Video condition x Half	0.59	0.553	0.00	0.17	0.845	0.00
Signal x Region x Video number x Video condition	0.22	0.801	0.00	0.43	0.652	0.00
Signal x Region x Video number x Half	0.05	0.954	0.00	0.21	0.814	0.00
Signal x Region x Video condition x Half	0.03	0.855	0.00	0.08	0.774	0.00
Signal x Video number x Video condition x Half	0.03	0.973	0.00	1.17	0.311	0.00
Region x Video number x Video condition x Half	0.11	0.896	0.00	0.52	0.593	0.00
Signal x Region x Video number x Video condition x Half	0.05	0.950	0.00	0.16	0.850	0.00

Table B 17

Mean, standard deviation and N for average relative power over video halves for test and retest sessions

Signal	*Region*	*Video number*	*Video condition*	*Video half*	TEST			RETEST		
					M	*SD*	*N*	*M*	*SD*	*N*
Alpha	Frontal	First	Non-social	First	0.31	0.09	29	0.28	0.07	37
				Second	0.30	0.08	29	0.28	0.08	37
			Social	First	0.30	0.07	27	0.28	0.06	37
				Second	0.30	0.08	27	0.30	0.08	37
		Second	Non-social	First	0.27	0.08	25	0.27	0.07	36
				Second	0.26	0.07	25	0.26	0.07	36
			Social	First	0.30	0.08	24	0.27	0.08	37
				Second	0.30	0.08	24	0.26	0.07	37
		Third	Non-social	First	0.28	0.08	25	0.26	0.08	34
				Second	0.28	0.07	25	0.26	0.09	34
			Social	First	0.29	0.08	26	0.26	0.08	34
				Second	0.28	0.08	26	0.25	0.07	34
	Posterior	First	Non-social	First	0.30	0.06	29	0.28	0.06	37
				Second	0.29	0.07	29	0.27	0.06	37
			Social	First	0.30	0.06	27	0.28	0.06	37
				Second	0.30	0.06	27	0.30	0.07	37
		Second	Non-social	First	0.26	0.07	25	0.26	0.06	36

				Second	0.27	0.07	25	0.26	0.07	36
			Social	First	0.30	0.08	24	0.27	0.07	37
				Second	0.30	0.08	24	0.28	0.07	37
		Third	Non-social	First	0.27	0.06	25	0.25	0.06	34
				Second	0.27	0.06	25	0.26	0.06	34
			Social	First	0.29	0.07	26	0.27	0.07	34
				Second	0.28	0.07	26	0.27	0.06	34
Theta	Frontal	First	Non-social	First	0.41	0.07	29	0.43	0.07	37
				Second	0.41	0.07	29	0.45	0.06	37
			Social	First	0.41	0.06	27	0.45	0.07	33
				Second	0.44	0.08	27	0.44	0.08	37
		Second	Non-social	First	0.43	0.08	25	0.42	0.07	36
				Second	0.43	0.06	25	0.42	0.08	36
			Social	First	0.43	0.09	24	0.44	0.08	37
				Second	0.44	0.08	24	0.45	0.08	37
		Third	Social	First	0.43	0.07	25	0.42	0.05	34
				Second	0.44	0.07	25	0.43	0.06	34
			Non-social	First	0.43	0.06	26	0.44	0.07	34
				Second	0.44	0.06	26	0.44	0.06	34
	Posterior	First	Non-social	First	0.37	0.06	29	0.39	0.05	37
				Second	0.38	0.07	29	0.39	0.06	37
			Social	First	0.37	0.06	27	0.41	0.06	37

		Second	0.40	0.07	27	0.42	0.07	37
Second	Non-social	First	0.38	0.06	25	0.38	0.06	36
		Second	0.39	0.06	25	0.40	0.07	36
	Social	First	0.39	0.06	24	0.41	0.07	37
		Second	0.42	0.07	24	0.41	0.07	37
Third	Non-social	First	0.38	0.06	25	0.38	0.06	34
		Second	0.39	0.06	25	0.39	0.06	34
	Social	First	0.40	0.06	26	0.40	0.07	34
		Second	0.41	0.07	26	0.42	0.06	34

Table B 18

*Estimate, t-value, p-value, and standard errors for pairwise comparisons between relative alpha and theta power in the first and second half of video viewing in the test session. Significant effects are indicated with a **

		estimate	*t*	*p*	*SE*
Alpha, first half	Alpha, second half	-0.001	-0.23	1.00	0.005
Theta, first half	Theta, second half	-0.007	-1.40	.975	0.005
Alpha, first half	Theta, first half	-0.146	-31.35	< .0001*	0.005
Alpha, second half	Theta, second half	-0.152	-32.60	< .0001*	0.005

Table B 19

Mean, standard error, and 95% confidence intervals for relative alpha and theta power in the first and second half of video viewing in the test session

Signal	Video condition	M	SE	95% CIs	
Alpha	First half	0.27	0.00	0.26	0.28
	Second half	0.27	0.00	0.26	0.28
Theta	First half	0.42	0.00	0.41	0.42
	Second half	0.42	0.00	0.41	0.43

Appendix B.3 Growth curves

Appendix B.3a Absolute

Linear mixed models to assess power change over the course of video viewing were rerun using a minimum of 30 usable EEG segments and including reference location as a random effect. For two models, frontal theta test and retest, including location reference as a random effect resulted in a poor model fit. For these models, this was removed and location reference was instead included as a fixed effect, to double check its impact. These models indicated no significant effect of reference location on power, further verifying that this did not impact findings.

Table B 20

Mean, standard error and 95% confidence intervals for first, second and third videos, and social and non-social conditions from average absolute power per segment in the test and retest session

Signal		TEST				RETEST			
		M	*SE*	95% CIs		*M*	*SE*	95% CIs	
				Video number					
Frontal alpha	First	1.69	0.11	1.48	1.91	1.62	0.1	1.43	1.81
	Second	1.59	0.11	1.38	1.81	1.58	0.09	1.39	1.77

	Third	1.61	0.11	1.40	1.83	1.55	0.09	1.37	1.74
				Video condition					
	Non-social	1.62	0.11	1.41	1.84	1.58	0.09	1.39	1.76
	Social	1.65	0.11	1.43	1.86	1.59	0.09	1.41	1.78
				Video number					
	First	1.9	0.13	1.64	2.17	1.94	0.09	1.77	2.11
	Second	1.84	0.13	1.58	2.1	1.91	0.09	1.74	2.08
Posterior alpha	Third	1.85	0.13	1.59	2.12	1.89	0.09	1.72	2.06
				Video condition					
	Non-social	1.84	0.13	1.57	2.1	1.89	0.09	1.72	2.06
	Social	1.89	0.13	1.63	2.16	1.94	0.09	1.77	2.11
				Video number					
	First	2.5	0.05	2.4	2.59	2.52	0.05	2.42	2.63
	Second	2.51	0.05	2.41	2.6	2.53	0.05	2.43	2.64
Frontal theta	Third	2.55	0.05	2.45	2.64	2.52	0.05	2.41	2.63
				Video condition					
	Non-social	2.5	0.05	2.41	2.6	2.49	0.05	2.39	2.6
	Social	2.53	0.05	2.44	2.62	2.56	0.05	2.45	2.66
				Video number					
	First	2.54	0.1	2.35	2.74	2.67	0.06	2.55	2.79
Posterior theta	Second	2.57	0.1	2.38	2.77	2.69	0.06	2.57	2.81
	Third	2.59	0.1	2.4	2.79	2.69	0.06	2.57	2.81

Video condition								
Non-social	2.53	0.1	2.33	2.72	2.63	0.06	2.51	2.75
Social	2.61	0.1	2.42	2.81	2.73	0.06	2.62	2.85

Table B 21

*Chi-squared, p-value and partial eta squared for segment number, video number and condition effects from an ANOVA of linear mixed models of average absolute power per segment number, video number and condition for frontal and posterior alpha and theta for the test and retest sessions. Significant effects are indicated with a **

		TEST		RETEST	
		χ2	*p*	*χ2*	*p*
Frontal alpha	Segment number	5.64	0.018	0.06	0.806
	Video number	33.75	<0.001*	13.74	0.001*
	Condition	2.81	0.094	0.85	0.358
Posterior alpha	Segment number	0.79	0.374	6.32	0.012*
	Video number	19.33	<0.001*	10.68	0.005*
	Condition	19.57	<0.001*	16.98	<0.001*
Frontal theta	Segment number	25.79	<0.001*	5.33	0.021*
	Video number	8.02	0.018*	0.53	0.768
	Condition	2.53	0.112	15.11	<0.001*
Posterior theta	Segment number	27.39	<0.001*	7.97	0.005*
	Video number	9.75	0.008*	2.67	0.263
	Condition	44.48	<0.001*	58.41	<0.001*

Table B 22

Estimate, standard error, degrees of freedom and 95% confidence intervals for each of the fixed effects in linear mixed models performed on average absolute power per segment for each of frontal and posterior alpha and theta in the test and retest sessions

	Fixed Effects									
	TEST					RETEST				
	β	*SE*	*DF*	CIs 95%		*β*	*SE*	*DF*	CIs 95%	
Frontal alpha										
Intercept	1.65	0.11	1.89	1.49	1.85	1.61	0.10	3.04	1.43	1.78
Segment Number	0.00	0.00	9436.00	0.00	0.00	0.00	0.00	8943.38	0.00	0.00
Video number: second	-0.10	0.02	9468.81	-0.13	-0.08	-0.04	0.02	8961.64	-0.06	-0.01
Video number: third	-0.08	0.02	9453.64	-0.10	-0.08	-0.07	0.02	8950.21	-0.09	-0.04
Video condition: social	0.02	0.01	9470.18	0.01	0.04	0.01	0.01	8959.12	-0.01	0.03
Posterior alpha										
Intercept	1.87	0.14	1.60	1.67	2.11	1.89	0.09	2.25	1.72	2.04
Segment Number	0.00	0.00	9435.69	0.00	0.00	0.00	0.00	8943.29	0.00	0.00
Video number: second	-0.07	0.02	9458.29	-0.09	-0.05	-0.03	0.02	8962.79	-0.04	0.00
Video number: third	-0.05	0.02	9465.52	-0.07	-0.05	-0.05	0.02	8954.57	-0.07	-0.03
Video condition: social	0.05	0.01	9459.69	0.04	0.07	0.05	0.01	8960.12	0.03	0.07
Frontal theta										
Intercept	2.42	0.05	60.79	2.37	2.53	2.46	0.06	49.27	2.35	2.54
Segment Number	0.00	0.00	9437.14	0.00	0.00	0.00	0.00	8943.85	0.00	0.00
Video number: second	0.01	0.02	9477.86	-0.02	0.03	0.01	0.02	8968.70	-0.01	0.05
Video number: third	0.05	0.02	9479.41	0.03	0.05	0.00	0.02	8960.18	-0.02	0.02
Video condition: social	0.02	0.02	9479.11	0.01	0.04	0.06	0.02	8966.16	0.03	0.09

Posterior theta										
Intercept	2.44	0.10	2.35	2.29	2.62	2.59	0.06	10.05	2.48	2.68
Segment Number	0.00	0.00	9435.85	0.00	0.00	0.00	0.00	8943.41	0.00	0.00
Video number: second	0.03	0.02	9462.43	0.01	0.05	0.02	0.02	8951.54	0.01	0.05
Video number: third	0.05	0.02	9442.46	0.03	0.05	0.02	0.02	8828.25	0.01	0.05
Video condition: social	0.09	0.01	9463.89	0.07	0.10	0.10	0.01	8961.58	0.08	0.12

Table B 23

Estimate and standard error for each of the random effects in linear mixed models performed on average absolute power per segment for each of frontal and posterior alpha and theta in the test and retest sessions

Random Effects					
		TEST		RETEST	
		β	*SE*	β	*SE*
Frontal alpha	ID: Intercept	0.13	0.37	0.14	0.38
	Reference location	0.02	0.13	0.01	0.10
	Residual	0.47	0.69	0.48	0.69
Posterior alpha	ID: Intercept	0.16	0.40	0.09	0.30
	Reference location	0.03	0.17	0.01	0.10
	Residual	0.34	0.58	0.33	0.58
Frontal theta	ID: Intercept	0.09	0.30	0.11	0.33
	Reference location	0.00	0.00	0.00	0.00

	Residual	0.54	0.73	0.54	0.73
Posterior theta	ID: Intercept	0.15	0.38	0.10	0.32
	Reference location	0.01	0.11	0.00	0.04
	Residual	0.38	0.61	0.38	0.62

Appendix B.3b Relative

Linear mixed models to assess power change over the course of video viewing were rerun using a minimum of 30 usable EEG segments and including reference location as a random effect. For both posterior alpha and frontal theta retest models the significance of the effect of segment number reduced, but did remain significant at the p <.05 level. For the model for posterior theta test, including location reference as a random effect resulted in a poor model fit. Location reference was therefore removed as a random effect and was instead included as a fixed effect, to double check its impact. These models indicated no significant effect of reference location on power, further verifying that this did not impact findings

Table B 24

Mean, standard error and 95% confidence intervals for first, second and third videos, and social and non-social conditions from average relative power per segment in the test and retest session

Signal		TEST				RETEST			
		M	*SE*	95% CIs		*M*	*SE*	95% CIs	
				Video number					
Frontal alpha	First	0.29	0.02	0.26	0.33	0.28	0.02	0.25	0.31
	Second	0.27	0.02	0.23	0.30	0.26	0.02	0.23	0.29
	Third	0.26	0.02	0.23	0.30	0.26	0.02	0.23	0.29

				Video condition					
	Non-social	0.27	0.02	0.23	0.31	0.26	0.02	0.23	0.29
	Social	0.28	0.02	0.24	0.31	0.27	0.02	0.24	0.30
				Video number					
	First	0.29	0.01	0.26	0.32	0.27	0.01	0.25	0.30
	Second	0.27	0.01	0.24	0.29	0.26	0.01	0.23	0.29
Posterior alpha	Third	0.26	0.01	0.24	0.29	0.26	0.01	0.23	0.28
				Video condition					
	Non-social	0.27	0.01	0.24	0.29	0.26	0.01	0.23	0.29
	Social	0.28	0.01	0.25	0.31	0.27	0.01	0.24	0.30
				Video number					
	First	0.43	0.01	0.41	0.46	0.44	0.01	0.42	0.46
	Second	0.43	0.01	0.41	0.46	0.43	0.01	0.41	0.45
Frontal theta	Third	0.44	0.01	0.41	0.46	0.43	0.01	0.41	0.45
				Video condition					
	Non-social	0.43	0.01	0.40	0.45	0.42	0.01	0.40	0.44
	Social	0.44	0.01	0.41	0.46	0.45	0.01	0.43	0.46
				Video number					
	First	0.39	0.01	0.37	0.40	0.40	0.01	0.37	0.42
Posterior theta	Second	0.39	0.01	0.38	0.41	0.39	0.01	0.37	0.42
	Third	0.39	0.01	0.37	0.40	0.39	0.01	0.37	0.42
				Video condition					

Non-social	0.38	0.01	0.36	0.39	0.38	0.01	0.36	0.40
Social	0.40	0.01	0.38	0.41	0.41	0.01	0.38	0.43

Table B 25

*Chi-squared, p-value and partial eta squared for segment number, video number and condition effects from an ANOVA of linear mixed models of average relative power per segment number, video number and condition for frontal and posterior alpha and theta for the test and retest sessions. Significant effects are indicated with a **

		TEST		RETEST	
		χ2	*p*	*χ2*	*p*
Frontal alpha	Segment number	0.68	0.410	0.95	0.330
	Video number	61.26	<0.001*	26.31	<0.001*
	Condition	3.39	0.066	1.22	0.269
Posterior alpha	Segment number	1.10	0.294	4.17	0.041*
	Video number	68.60	<0.001*	29.03	<0.001*
	Condition	25.72	<0.001*	21.47	<0.001*
Frontal theta	Segment number	5.17	0.023*	3.88	0.049
	Video number	2.15	0.341	3.25	0.197
	Condition	5.36	0.021*	36.98	<0.001*
Posterior theta	Segment number	22.14	<0.001*	9.75	0.002*
	Video number	2.90	0.235	2.30	0.316
	Condition	66.05	<0.001*	88.81	<0.001*

Table B 26

Estimate, standard error, degrees of freedom and 95% confidence intervals for each of the fixed effects in linear mixed models performed on average relative power per segment for each of frontal and posterior alpha and theta in the test and retest sessions

Fixed Effects

	TEST					RETEST				
	β	*SE*	*DF*	CIs 95%		*β*	*SE*	*DF*	CIs 95%	
				Frontal alpha						
Intercept	0.29	0.02	2.52	0.26	0.33	0.28	0.02	3.58	0.25	0.31
Segment Number	0.00	0.00	9436.45	0.00	0.00	0.00	0.00	8944.39	0.00	0.00
Video number: second	-0.03	0.00	9472.81	-0.03	-0.02	-0.02	0.00	8967.72	-0.02	-0.01
Video number: third	-0.03	0.00	9413.25	-0.03	-0.03	-0.02	0.00	8926.45	-0.02	-0.01
Video condition: social	0.01	0.00	9474.29	0.00	0.01	0.00	0.00	8966.09	0.00	0.01
				Posterior alpha						
Intercept	0.28	0.01	3.71	0.26	0.31	0.26	0.01	2.71	0.24	0.29
Segment Number	0.00	0.00	9436.28	0.00	0.00	0.00	0.00	8944.07	0.00	0.00
Video number: second	-0.02	0.00	9470.79	-0.03	-0.02	-0.01	0.00	8967.90	-0.01	-0.01
Video number: third	-0.03	0.00	9324.31	-0.03	-0.03	-0.02	0.00	8945.68	-0.02	-0.01
Video condition: social	0.01	0.00	9472.68	0.01	0.02	0.01	0.00	8966.42	0.01	0.02
				Frontal theta						
Intercept	0.42	0.01	3.27	0.40	0.44	0.42	0.01	14.17	0.40	0.44
Segment Number	0.00	0.00	9438.48	0.00	0.00	0.00	0.00	8913.74	0.00	0.00
Video number: second	0.00	0.00	9477.37	0.00	0.01	-0.01	0.00	8957.26	-0.01	0.00
Video number: third	0.01	0.00	8966.80	0.00	0.01	-0.01	0.00	6809.80	-0.01	0.00
Video condition: social	0.01	0.00	9477.48	0.00	0.01	0.02	0.00	8973.21	0.02	0.03
				Posterior theta						

Intercept	0.36	0.01	64.37	0.36	0.38	0.38	0.01	4.23	0.36	0.40
Segment Number	0.00	0.00	9438.03	0.00	0.00	0.00	0.00	8944.52	0.00	0.00
Video number: second	0.01	0.00	9479.39	0.00	0.01	0.00	0.00	8971.77	-0.01	0.00
Video number: third	0.00	0.00	9477.20	0.00	0.00	-0.01	0.00	8879.58	-0.01	0.00
Video condition: social	0.02	0.00	9479.95	0.02	0.03	0.03	0.00	8969.91	0.02	0.03

Table B 27

Estimate and standard error for each of the random effects in linear mixed models performed on average relative power per segment for each of frontal and posterior alpha and theta in the test and retest sessions

Random Effects					
		TEST		RETEST	
		β	*SE*	β	*SE*
Frontal alpha	ID: Intercept	0.00	0.07	0.00	0.06
	Reference location	0.00	0.02	0.00	0.02
	Residual	0.02	0.15	0.02	0.14
Posterior alpha	ID: Intercept	0.00	0.06	0.00	0.05
	Reference location	0.00	0.01	0.00	0.02
	Residual	0.02	0.12	0.01	0.12
Frontal theta	ID: Intercept	0.00	0.05	0.00	0.06
	Reference location	0.00	0.01	0.00	0.00
	Residual	0.03	0.16	0.03	0.16

Posterior theta	ID: Intercept	0.00	0.05	0.00	0.05
	Reference location	0.00	0.00	0.00	0.01
	Residual	0.02	0.13	0.02	0.13

Appendix B.4 Test-retest reliabilities

Table B 28

ICC statistics for reliability analyses between test and retest sessions for a number of measures. Statistics are coloured according to ICC values: values below 0.40 are red, values from 0.40 to 0.75 are orange, and values above 0.75 are green

		Absolute										Relative									
	Condition	Alpha					Theta					Alpha					Theta				
		ICC	LB	UB	*p*	*n*	ICC	LB	UB	*p*	*n*	ICC	LB	UB	*p*	*n*	ICC	LB	UB	*p*	*n*
Measure	Average across regions																				
	Social, first	0.362	-0.03	0.66	0.034	25	0.740	0.49	0.88	0.000	25	0.503	0.15	0.74	0.004	25	0.728	0.48	0.87	0.000	25
	Social, second	0.507	0.18	0.73	0.002	29	0.597	0.30	0.79	0.000	29	0.544	0.23	0.76	0.001	29	0.785	0.59	0.89	0.000	29
	Social, third	0.511	0.15	0.75	0.004	25	0.736	0.49	0.87	0.000	25	0.608	0.29	0.81	0.001	25	0.718	0.44	0.87	0.000	25
	Nonsocial, first	0.613	0.29	0.81	0.000	23	0.688	0.39	0.86	0.000	23	0.750	0.47	0.89	0.000	23	0.796	0.54	0.91	0.000	23
	Nonsocial, second	0.672	0.36	0.85	0.000	23	0.773	0.54	0.90	0.000	23	0.680	0.38	0.85	0.000	23	0.840	0.66	0.93	0.000	23
	Nonsocial, third	0.555	0.16	0.79	0.004	21	0.805	0.59	0.92	0.000	21	0.608	0.25	0.82	0.001	21	0.696	0.39	0.86	0.000	21
	Social	0.329	-0.60	0.85	0.227	7	0.867	0.40	0.98	0.002	7	0.757	0.09	0.95	0.017	7	0.867	0.47	0.98	0.001	7
	Nonsocial	0.693	-0.08	0.94	0.035	7	0.835	0.37	0.97	0.003	7	0.750	0.04	0.95	0.021	7	0.915	0.60	0.98	0.000	7
	First	0.428	-0.01	0.73	0.027	20	0.715	0.42	0.88	0.000	20	0.656	0.32	0.85	0.000	20	0.805	0.55	0.92	0.000	20
	Second	0.620	0.26	0.83	0.001	21	0.780	0.53	0.90	0.000	21	0.633	0.29	0.83	0.001	21	0.857	0.68	0.94	0.000	21
Average over whole video	Third	0.435	-0.07	0.76	0.043	16	0.860	0.65	0.95	0.000	16	0.636	0.23	0.86	0.003	16	0.793	0.51	0.92	0.000	16
	Difference between regions																				
	Social, first	0.600	0.27	0.80	0.001	25	0.424	0.04	0.70	0.016	25	0.289	-0.10	0.61	**0.070**	25	0.650	0.35	0.83	0.000	25
	Social, second	0.307	-0.07	0.60	0.054	29	0.347	-0.02	0.63	0.033	29	0.475	0.13	0.72	0.005	29	0.284	-0.09	0.59	**0.064**	29
	Social, third	0.639	0.33	0.82	0.000	25	0.397	0.00	0.68	0.024	25	0.466	0.09	0.73	0.010	25	0.123	-0.28	0.49	**0.274**	25
	Nonsocial, first	0.472	0.07	0.74	0.012	23	0.127	-0.29	0.50	**0.275**	23	0.207	-0.21	0.56	**0.161**	23	-0.086	-0.48	0.33	**0.653**	23
	Nonsocial, second	0.486	0.09	0.75	0.009	23	0.315	-0.12	0.64	**0.073**	23	0.490	0.10	0.75	0.008	23	0.444	0.05	0.72	0.014	23
	Nonsocial, third	-0.130	-0.50	0.32	0.722	18	-0.004	-0.50	0.47	**0.506**	18	0.012	-0.27	0.38	**0.470**	18	-0.110	-0.53	0.33	**0.682**	22
	Social	0.725	0.04	0.95	0.020	7	0.808	0.30	0.96	0.004	7	0.693	0.05	0.94	0.018	7	0.238	-0.75	0.82	**0.308**	7
	Nonsocial	0.649	0.00	0.93	0.025	7	0.567	-0.15	0.91	**0.054**	7	0.603	-0.09	0.92	0.041	7	0.264	-0.58	0.82	**0.265**	7
	First	0.579	0.19	0.81	0.004	20	0.269	-0.18	0.63	**0.117**	20	0.173	-0.24	0.55	**0.209**	20	0.352	-0.09	0.68	**0.058**	20
	Second	0.345	-0.11	0.67	0.064	21	0.312	-0.15	0.65	**0.086**	21	0.648	0.31	0.84	0.001	21	0.408	-0.03	0.71	0.034	21
Average over whole video	Third	0.735	0.39	0.90	0.000	16	0.563	0.10	0.82	0.010	16	0.424	-0.08	0.75	0.048	16	0.162	-0.38	0.60	**0.275**	16
	Average across regions																				
	Social, first	-0.555	-0.83	-0.14	**0.994**	22	0.260	-0.09	0.58	**0.079**	22	-0.503	-0.80	-0.08	**0.989**	22	-0.136	-0.41	0.23	**0.785**	22
	Social, second	0.371	-0.06	0.68	0.044	22	0.172	-0.16	0.51	**0.164**	22	0.428	0.02	0.71	0.021	23	0.001	-0.30	0.35	**0.498**	23
	Social, third	0.304	-0.15	0.65	**0.092**	21	0.133	-0.33	0.53	**0.284**	21	-0.174	-0.59	0.29	**0.768**	21	-0.183	-0.60	0.28	**0.780**	21
	Nonsocial, first	0.398	-0.04	0.71	0.037	21	-0.219	-0.57	0.21	**0.846**	21	0.456	0.06	0.73	0.013	21	0.068	-0.35	0.47	**0.379**	21
	Nonsocial, second	0.132	-0.39	0.57	**0.309**	17	0.461	-0.03	0.77	0.032	17	0.197	-0.33	0.62	**0.226**	17	-0.467	-0.83	0.05	**0.962**	17
	Nonsocial, third	-0.258	-0.69	0.25	**0.840**	18	0.301	-0.20	0.67	**0.114**	18	-0.282	-0.68	0.21	**0.870**	18	0.100	-0.35	0.50	**0.331**	22
	Social	-1.052	-2.19	0.98	0.826	2	-0.398	-0.56	0.84	**0.926**	2	-0.041	-0.40	0.99	**0.553**	2	-0.149	-0.16	-0.16	**0.996**	2
	Nonsocial	-1.859	-1.87	-1.87	1.000	2	-0.041	-6.42	1.00	**0.509**	2	-1.736	-1.77	-1.77	**0.999**	2	-22.485	-22.99	-22.99	**1.000**	2
	First	-0.050	-0.51	0.42	0.579	18	-0.071	-0.51	0.40	**0.613**	18	-0.435	-0.79	0.07	**0.957**	18	-0.175	-0.52	0.27	**0.790**	18
	Second	0.100	-0.57	0.65	0.385	11	0.452	-0.09	0.81	**0.052**	11	0.427	-0.21	0.80	**0.083**	11	-0.029	-0.59	0.54	**0.537**	12
Difference between halves of video	Third	0.287	-0.37	0.73	0.186	12	0.162	-0.49	0.67	**0.311**	12	0.272	-0.31	0.71	**0.176**	12	-0.127	-0.62	0.40	**0.676**	16
	Difference between regions																				
	Social, first	0.437	0.05	0.72	0.015	22	0.281	-0.17	0.63	**0.104**	22	0.373	-0.06	0.68	0.043	22	0.542	0.16	0.78	0.005	22
	Social, second	0.321	-0.10	0.65	0.064	22	-0.252	-0.58	0.17	**0.886**	22	0.015	-0.38	0.41	**0.472**	23	0.014	-0.32	0.38	**0.469**	23
	Social, third	-0.367	-0.72	0.09	0.944	21	0.431	0.04	0.72	0.017	21	0.233	-0.23	0.60	**0.156**	21	0.431	0.00	0.73	0.026	21
	Nonsocial, first	-0.065	-0.46	0.36	0.617	21	-0.117	-0.52	0.32	**0.698**	21	-0.256	-0.65	0.21	**0.861**	21	-0.032	-0.42	0.38	**0.560**	21
	Nonsocial, second	-0.263	-0.68	0.25	0.845	17	-0.113	-0.59	0.39	**0.664**	17	0.229	-0.30	0.64	**0.190**	17	-0.378	-0.78	0.15	**0.922**	17
	Nonsocial, third	-0.130	-0.50	0.32	0.722	18	-0.004	-0.50	0.47	**0.506**	18	0.012	-0.27	0.38	**0.470**	18	-0.110	-0.53	0.33	**0.682**	22
	Social	0.070	-0.09	0.99	0.308	2	0.846	-0.29	1.00	**0.060**	2	-0.690	-0.69	-0.69	**1.000**	2	0.969	0.16	1.00	0.016	2
	Nonsocial	0.408	-0.80	1.00	0.272	2	0.192	-2.38	1.00	**0.433**	2	0.194	0.00	0.99	**0.060**	2	0.657	-35.88	1.00	**0.330**	2
	First	-0.010	-0.43	0.43	0.517	18	0.029	-0.47	0.49	**0.456**	18	-0.354	-0.74	0.15	**0.919**	18	0.377	-0.11	0.71	**0.062**	18
	Second	0.398	-0.19	0.79	0.086	11	-0.251	-0.82	0.42	**0.761**	11	0.399	-0.24	0.78	**0.100**	12	-0.303	-0.76	0.32	**0.837**	12
Difference between halves of video	Third	-0.189	-0.69	0.41	0.731	12	0.571	0.07	0.85	0.014	12	0.223	-0.28	0.67	**0.199**	12	0.458	-0.04	0.77	0.034	16
	LMER																				
	Differences	0.646	0.40	0.80	0.000	35	0.436	0.13	0.67	0.003	35	0.639	0.40	0.80	0.000	35	0.621	0.37	0.79	0.000	35
	Wholehead	0.677	0.45	0.82	0.000	35	0.484	0.19	0.70	0.001	35	0.738	0.54	0.86	0.000	35	0.734	0.54	0.86	0.000	35
	Frontal	0.663	0.43	0.81	0.000	35	0.451	0.15	0.68	0.002	35	0.690	0.47	0.83	0.000	35	0.678	0.45	0.82	0.000	35
Intercepts from LMERs	Posterior	0.673	0.43	0.82	0.000	35	0.666	0.38	0.83	0.000	35	0.732	0.53	0.85	0.000	35	0.773	0.57	0.88	0.000	35
	Variability																				
	Differences	-0.089	-0.38	0.23	0.710	35	0.122	-0.22	0.44	**0.243**	35	0.323	-0.01	0.59	0.030	35	-0.194	-0.51	0.15	**0.865**	35
	Wholehead	0.433	0.12	0.67	0.005	35	0.277	-0.06	0.56	**0.054**	35	0.420	0.10	0.66	0.006	35	0.449	0.14	0.68	0.003	35
	Frontal	0.340	0.02	0.60	0.019	35	0.366	0.04	0.62	0.015	35	0.458	0.15	0.68	0.003	35	0.282	-0.06	0.56	**0.051**	35
RSE	Posterior	0.295	-0.04	0.57	0.041	35	0.041	-0.30	0.37	**0.408**	35	0.347	0.01	0.61	0.021	35	0.290	-0.05	0.57	0.046	35
	Differences	-0.100	-0.39	0.22	0.731	35	0.025	-0.32	0.36	**0.443**	35	0.263	-0.08	0.55	**0.064**	35	-0.229	-0.53	0.12	**0.904**	35
	Wholehead	0.398	0.07	0.65	0.009	35	0.289	-0.05	0.57	0.046	35	0.465	0.16	0.69	0.002	35	0.481	0.18	0.70	0.002	35
	Frontal	0.322	0.00	0.59	0.026	35	0.361	0.03	0.62	0.016	35	0.516	0.22	0.72	0.001	35	0.305	-0.03	0.58	0.037	35
SD	Posterior	0.269	-0.07	0.55	0.056	35	0.073	-0.27	0.40	**0.339**	35	0.376	0.05	0.63	0.013	35	0.342	0.01	0.61	0.022	35
	Differences	0.382	0.06	0.63	0.011	35	0.414	0.09	0.66	0.007	35	0.337	0.03	0.59	0.016	35	0.396	0.08	0.64	0.008	35
	Wholehead	0.353	0.03	0.61	0.015	35	0.359	0.04	0.61	0.015	35	0.352	0.04	0.61	0.015	35	0.370	0.05	0.62	0.013	35
	Frontal	0.359	0.04	0.61	0.015	35	0.365	0.04	0.62	0.013	35	0.351	0.04	0.61	0.015	35	0.373	0.05	0.62	0.012	35
SE_Ranef	Posterior	0.348	0.03	0.60	0.016	35	0.354	0.03	0.61	0.015	35	0.353	0.04	0.61	0.015	35	0.368	0.05	0.62	0.013	35

APPENDIX C. QUESTIONNAIRES AS REPORTED IN CHAPTER 4

Appendix C.1 Home CHAOS questionnaire

For each statement below, please assign a number between 1 and 4 to indicate how much each statement describes your home environment. Please use the following scale:

1. There is very little commotion in our home.
2. We can usually find things when we need them.
3. We almost always seem to be rushed.
4. We are usually able to stay on top of things.
5. No matter how hard we try, we always seem to be running late.
6. It's a real zoo in our home.
7. At home we can talk to each other without being interrupted.
8. There is often a fuss going on at our home.
9. No matter what our family plans, it usually doesn't seem to work out.
10. You can't hear yourself think in our home.
11. I often get drawn into other people's arguments at home.
12. Our home is a good place to relax.
13. The telephone takes up a lot of our time at home.
14. The atmosphere in our home is calm.
15. First thing in the day, we have a regular routine at home.

Appendix C.2 Neighbourhood Safety Scale

	Perceptions of neighborhood conditions (Based on responses to the following survey questions)
Crime and Safety	How often these things are a problem or are found in your neighborhood?/How worried are you about the following things in your neighborhood: *Range: 1(Rarely/Not worried) to 10 (Frequently/Very worried)* (1) Drug Dealers or users hanging around (2) Having property stolen (3) Walking alone during the day (4) Letting children go outside during the day (5) Letting children go outside during the night (6) Being robbed (7) Being murdered
Physical Disorder	How often these things are a problem or are found in your neighborhood? *Range: 1 (Rarely) to 10 (Frequently)* (1) Litter or trash on the sidewalks or streets (2) Graffiti on buildings and walls (3) Abandoned cars (4) Vacant, abandoned or boarded up buildings (5) Houses and yards not kept up
Social Disorder	How often are these things a problem or are found in your neighborhood? *Range: 1 (Rarely) to 10 (Frequently)* (1) Drunks hanging around (2) Unemployed adults hanging around (3) Young adults hanging around (4) Gang activity

Appendix C.3 GAD-7 Anxiety Scale

GAD-7 Anxiety

Over the last two weeks, how often have you been bothered by the following problems?	Not at all	Several days	More than half the days	Nearly every day
1. Feeling nervous, anxious, or on edge	0	1	2	3
2. Not being able to stop or control worrying	0	1	2	3
3. Worrying too much about different things	0	1	2	3
4. Trouble relaxing	0	1	2	3
5. Being so restless that it is hard to sit still	0	1	2	3
6. Becoming easily annoyed or irritable	0	1	2	3
7. Feeling afraid, as if something awful might happen	0	1	2	3

Appendix C.4 BIS Anxiety Scale

Your child usually gets very tense when she or he thinks something unpleasant will happen.

Your child worries about making mistakes.

Your child is hurt when people scold him or her or tell that she or he does something wrong.

Your child feels pretty upset when she or he thinks that someone is angry with him or her.

Your child does not become fearful or nervous, even when something bad happens to him or her (R).

Your child feels worried when she or he thinks she or he has done poorly at something.

Your child is very fearful compared to his or her friends.

Appendix C.5 Parent Feedback Form questions

1. Please rate how satisfied you were with the location of the visit?

Not Satisfied Fairly Satisfied Very Satisfied

1	2	3	4	5

2. Please rate how satisfied you were with the duration of the visit?

Not Satisfied Fairly Satisfied Very Satisfied

1	2	3	4	5

3. Please rate how satisfied you were with the assessment of your child?

Not Satisfied Fairly Satisfied Very Satisfied

1	2	3	4	5

4. Please rate how satisfied you were with the way the study and reasons for the study were explained to you and your family?

Not Satisfied Fairly Satisfied Very Satisfied

1	2	3	4	5

5. Please rate how important you feel this study is according to the aim as outlined on the study information sheet

No importance　　　　Fairly important　　　　Very important

1	2	3	4	5

6. Please rate whether you think this assessment is a good way to assess children's neurodevelopment from a parent's point of view?

Not useful　　　　Fairly useful　　　　Very useful

1	2	3	4	5

APPENDIX D. QUESTIONNAIRES USED IN CHAPTER 5

Appendix D.1 Focus group study: Questionnaire (a) about participants' views of research

Questions about your view of research

- What do you understand by the term 'research'?

____

____

____

- What kinds of research do you consider most important? Please rank the following areas with most important as number 1 and least important as 9

	Economic
	Medicine
	Engineering (Robotics)
	Space
	Chemistry
	Climate
	Education
	Genetics
	Sports and active living

- What level of research experience do you have?

	None
	A little (i.e. during school)
	A little more (i.e. some at university or post school, but limited)
	Some (i.e. conducted own research at University)
	Lots (have conducted more than one study)
	Extensive (have multiple years of research experience, have published papers, PhD level, etc.)

- Have you taken part in any research before today? (please circle)

Yes	No

- If yes, how many studies?

______

Appendix D.2 Focus group study: Questionnaire (b) about participants' views of the factors influencing child development

Questions about your view of child development

- What do you understand by the term 'child development'?

____

____

____

- How interested are you in learning more about child development? Please place a cross on the line.

____

Not at all interested | Very interested

- What aspects of child development are you most interested in or would you most like to learn more about? Please rank the following areas with most important as number 1 and least important as 11.

	Emotional development
	Physical development
	Communication and language development
	Literacy
	Maths
	Being creative and imaginative
	Understanding of the World
	Problem solving
	Attention
	Memory
	Executive functioning

- There are many factors which can impact children's development. Which of the following would you most like to know more about? Please rank the following areas with the one you would most like to know more about as number 1.

	Sleep
	Diet / Nutrition
	Genetics (i.e. what children inherit from their parents)
	Gender or sex
	Exercise
	Parenting
	Social/ cultural practices (i.e. how different cultures communicate, educate, etc.)
	Education
	Family health/ stress
	Parental education, income, employment
	Siblings

Appendix D.3 Focus group study: Questionnaire (c) about the iTapp app and what parents would like from an app-based tool

Questions about the app

- How easy was the app to use?

Very difficult	
Quite difficult	
Neither easy nor difficult	
Quite easy	
Very easy	

- Which aspects of the app would you improve and how? (i.e. how it looks, the activities, how it functions, etc.). Please be honest and provide as many pointers as you'd like!

____

____

____

____

- How often would you use an app like this?

Never	
Rarely	
Once every few months	
Once a month	
Once a week	
Multiple times a week	
Once a day	
Multiple times a day	

- What factors would make you more or less likely to use this app? Please number the most important factors with 1 as most important.

	Easy to use
	Functioning of the app (i.e. if it has glitches, etc.)
	Appearance
	Quality of information it provides
	Quality of activities
	Entertainment/ learning opportunities for my child
	Founded in science, provides link to child development researchers
	Other, please specify:

- How clear was the information in the app?

Not at all clear	
A little unclear	
Neither clear nor unclear	
Quite clear	
Very clear	

- What would be your primary goal of using this app (i.e. for fun, to learn, etc.)?

- What would you like to get from this app? Please rank the following with 1 as the thing you would most like from the app.

	Way to track my child's development
	Provide snippets of scientific information about my child's particular stage of development
	Way to communicate with child development researchers
	Links to further resources about my child's development
	Information about how my child is performing in the activities
	Entertainment for my child
	Information about my child's development compared to other children (this would be compared to a big number of children, not individuals)
	Other, please provide details:

- What personal information would you be happy to give through the app (all information would be kept anonymous)? Please tick any you'd be happy to give.

	Parental education
	Size of house
	Family size
	Employment
	Income
	Race/ Ethnicity
	Medical history
	Information about resources in home (i.e. number of books, games, etc.)

Appendix D.4 Teachbrite app study: Sociodemographic questionnaire

We would like to collect demographic information on you and your family. This helps us to understand your family circumstances and helps us to make sure we include children in the study who are from a variety of backgrounds. We know that this can be sensitive information, so if you don't want to answer the questions feel free to select the 'do not wish to answer' option. Thank you!

1. **How many people (not including you) live in your home?**

2. **Of these, how many are children (under 18)?**

3. **How many bedrooms do you have in your home?**

4. **Which of the following best describes the highest level of education you have completed?**
    a. Didn't finish secondary school
    b. Didn't finish secondary school, but completed a technical/ vocational program
    c. Secondary school – finished school but fewer than 4 passes in GCSEs, O levels or equivalent
    d. Secondary school – passed at least 4 GCSEs, O levels or equivalent
    e. Further education – completed A-levels, BTECs or equivalent
    f. Higher education – undergraduate level (undergraduate degree, diploma, NVQ, etc)
    g. Higher education – postgraduate level (masters, doctorate, PGCE, etc)
    h. Other, please specify ________
    i. Do not wish to answer

5. **What is the highest year of education you have completed? *(Tick one)***

Primary School	Secondary School	College or Sixth Form	Undergraduate Degree	Postgraduate Degree
_____Reception	_____Year 7	_____Year 12	_____Year 1	_____Year 1
_____Year 1	_____Year 8	_____Year 13	_____Year 2	_____Year 2
_____Year 2	_____Year 9		_____Year 3	_____Year 3
_____Year 3	_____Year 10		_____Year 4	_____Year 4+
_____Year 4	_____Year 11			
_____Year 5				
_____Year 6				

6. **Do you have firm plans for further education?**
    a. If yes, what? ___

7. **What is your current employment status?**
    a. Disabled (not working due to permanent or temporary disability)
    b. Sick leave?
    c. Homemaker
    d. Retired
    e. Not currently employed, looking for work
    f. Working part time → number of hours per week ___
    g. Working full time
    h. Do not wish to answer

7b. If you are working, what kind of work do you do?

(Job Title)

____

8. **What best describes your yearly household income (before tax)? If you are not sure, please go with your best estimate.**

- <£20,00
- £20,000 - £29,999
- £30,000 - £39,999
- £40,000 - £59,999
- £60,000 - £79,999
- £80,000 - £99,999
- £100,000 - £149,999
- > £149,999
- Don’t know or do not wish to answer

9. **Instructions: Think of this ladder as representing where people stand in the UK.**
   At the **top** of the ladder are the people who are the best off – those who have the most money, the most education, and the most respected jobs. At the **bottom** are the people who are the worst off – those who have the least money, least education, the least respected jobs, or no job. The higher up you are on this ladder, the closer you are to the people at the very top; the lower you are, the closer you are to the people at the very bottom.

   Where would you place yourself on this ladder?

   Please place a large **“X”** on the rung where you think you stand at this time in your life relative to other people in the UK.

10. Instructions: Think of this ladder as representing where people stand in their communities.

People define community in different ways; please define it in whatever way is most meaningful to you. At the **top** of the ladder are people who have the highest standing in their community. At the **bottom** are the people who have the lowest standing in their community.

Where would you place yourself on this ladder?

Please place a large **"X"** on the rung where you think you stand at this time in your life relative to other people in your community.

Sources:

Adler, N. E., Epel, E. S., Castellazzo, G., & Ickovics, J. R. (2000). Relationship of subjective and objective social status with psychological and physiological functioning: Preliminary data in healthy, White women. *Health Psychology, 19*(6), 586-592.

MacArthur Scale of Subjective Social Status - Adult Version