Interactive brains:

How infant cognition interacts with the dynamic social world

Sayaka Kidby, MPhil.

Submitted in partial fulfilment of the requirement
for the degree of Doctor of Philosophy

Department of Psychology
Lancaster University

December 2021

TABLE OF CONTENTS

Declaration

I declare that this thesis is my own work completed under the supervision of Vincent Reid and Kirsty Dunn. Author contributions are listed at the end of each chapter. No portion of the work referred to in this thesis has been submitted elsewhere in support of application for another degree or qualification at this or any other institution.

Sayaka Kidby
24 December 2021

Thesis abstract

Research taking a cognitive neuroscience approach has shed light on social cognition during infancy. These studies have provided invaluable knowledge about how infants process social information, but a number of concepts regarding infant social cognition are often discussed based on research utilising rigidly controlled experimental paradigms where the role of infants is typically passive as an observer of stimuli. Increasing evidence suggests differences between the social cognitive processes that occurs when we act as observers of others (a ‘third-person’ perspective) and the processes that emerge when we are actively engaging with other people in an interactional context (a ‘second-person’ perspective) (e.g., Redcay and Schilbach, 2019; Siposova & Carpenter, 2019). Accordingly, there has been a growing recognition that we need a ‘second-person’ perspective, as compared to conventional “third-person” approach.

The aim of the current thesis is to explore the interplay between infant cognition and the social world surrounding them, by moving research settings to a more naturalistic and dynamic one where infants are positioned as part of interaction. Towards this goal, Study 1 (Chapter 2) reviewed the current progress of “second-person” neuroscience research to evaluate the validity and robustness of simultaneous dual brain scanning techniques, often referred to as hyperscanning. The review identified large heterogeneity in reported effect sizes between published studies, suggesting the need to improve comparability of research, such as establishing standardised methods or promoting open science practices including code and data sharing to achieve higher reproducibility. This thesis then turned to research using various techniques from a conventional screen-based paradigm to a more dynamic setting, with the aim of building a stable platform towards second-person cognitive neuroscience approaches that investigate infant cognition while the infant actively interacts with other people. Study 2 (Chapter 3) explored how infants encode information differently from two adults who give gaze cues to a target object with different levels of accuracy. Whilst the study utilised a conventional event-related potential paradigm using screen-based stimuli, this paradigm could be adapted to enable future studies to investigate how infants’ social cognitive ability to discriminate reliable and unreliable informants can inform their subsequent behaviour observed in a social interactional behavioural task. Study 3 (Chapter 4) moved towards the use of more dynamic video stimuli and explored the neural processing of unexpected events. The study identified challenges in using dynamic perceptual inputs as stimuli. Study 4

(Chapter 5) transitioned into more naturalistic social contexts and analysed infant cognition while 10-month-old infants were faced with an adult demonstrating novel object labels in a live interaction. The study not only showed the feasibility of second-person neuroscientific research with infant participants, but also advanced our knowledge about infant word learning a step further, and demonstrated the trajectory from the encoding of semantic word information to its consolidation as knowledge. Study 5 (Chapter 6) also utilised a naturalistic interactional setting where infants were able to actively engage in a social task with an experimenter in a live manner, and aimed to identify systematic differences in neural activity between 9-month-old infants who make perseverative errors originally reported by Piaget (1954) and those who do not. This study was, to our knowledge, the first of its kind to validate the feasibility of utilising neurophysiological measures in this traditional interactive behavioural paradigm, in such a way that it does not interfere with the standard procedure.

This thesis produced a series of studies which jointly demonstrate the potential for conducting research in a more dynamic setting that investigates infant social cognition taking a ‘second-person’ cognitive neuroscience approach to advance our knowledge about the intricate interaction between infant cognition, behaviour and the environment. We conclude this thesis by addressing the challenges of such an approach, to which we also attempt to propose solutions, as well as discussing future directions for the field.

Acknowledgement

This thesis could not have been possible without the support from my supervisors, Professor Vincent Reid and Dr Kirsty Dunn. You two have made a great supervisory team, and I cannot thank you enough for your critical and insightful feedback on my work, continuous encouragement as well as patience. You introduced me to many new ways of thinking, and provided me with models to follow to become a better researcher. Completing my PhD was not easy, but you made it enjoyable and stimulating.

I thank the Lancaster Babylab team. A special thanks to Gert, Szilvi, Kat, Uschi and Barrie. Thank you for being always willing to help in any way throughout my PhD, professionally and emotionally. Chiara, you deserve a special mention. You have been the best PhD partner that I could have asked for. Thank you so much for your incredible kindness and compassion, and being a sounding board when I needed it.

I am very grateful to the European Commission Marie Skłodowska-Curie Actions for generously supporting my professional development over the past three years, and foremost, for the opportunities for friendship and collaboration across countries. Aude and Lisanne, thank you for being such an inspiration as colleagues and friends, and all the laughs we shared.

I am also thankful to my families, in Japan and England. Dan, I could not have done this without you. We both had ups and downs during the last few years, but we made it through because we had each other.

Last but not least, I must extend my appreciation to all the parents and children who took part in and helped with my research. Many people say it is challenging to work with tiny humans, and whilst it is true, that has been the best part of my PhD as a path to become a developmental scientist.

List of tables

List of figures

Chapter 1

Towards investigations of neural mechanisms underlying infant social cognition in more dynamic settings

Infants as active participants of a reciprocal social interaction

Human development happens in a social environment. As social beings, we spend the majority of our time being with or thinking about others, and much of our learning occurs through interactions with others (Frith & Frith, 2001; Tomasello, Kruger, & Ratner, 1993). The importance of social interaction is critical during infancy, as social connection is essential for the optimal development of cognitive and socioemotional skills as well as language (Hinde & Stevenson-Hinde, 1987; Nelson, Furtado, Fox, & Zeanah, 2009). Infancy is also a key time during which the major development of the neural systems occur which underpins the processing of social information (Frith & Frith, 2010; Ilyka, Johnson, & Lloyd-Fox, 2021). The development of the brain is a process that is influenced by the interplay between intrinsic factors, such as molecular and cellular systems, and extrinsic factors such as the surrounding social environment (Jablonka, 2015; Keller, 2000; Sarkart, 2000; Stiles, 2017). Synaptic connections in the brain are built and strengthened, or selectively eliminated, based on how frequently they are activated (Singer, 1995). This points to the importance of experiencing quality social interactions during infancy, as it enables infant brains to strengthen functions that support social interactions (Kolb, Harker, & Gibb, 2017). Foremost, understanding how infant cognition works and develops in a social situation is critical, as it could help achieve an optimal social environment for young children to enjoy positive developmental outcomes.

From early stages of life, infants contribute to creating the systems and processes of mutual regulation and coordination between themselves and their social partners (Field, 1978; Tronick, 1989). Infants engage with others through non-verbal signals, and these signals from infants inform their parents and caregivers of what the infants need and what the parents should provide them with (Cohn & Tronick, 1988; Fogel, Diamond, Langhorst, & Demos, 1982; Tronick, Heidelise, & Brazelton, 1980). Eye gaze is one of the most important communicative cues between adults and preverbal infants (Scaife & Bruner, 1975). Eye gaze conveys not only the visual attentional target of another person but also their willingness to communicate (Baron-Cohen, 1995; Csibra, 2010). Therefore, it is not surprising that human infants are sensitive to eyes and the direction of gaze from birth, and they can reliably follow other people's gaze to locate their attentional target placed nearby from as early as 3 months of age (Batki et al, 2000; D'Entremont, Hains, & Muir, 1997; Farroni et al, 2002). Over the course of the first postnatal year, infants' gaze following behaviour becomes more nuanced (Butterworth & Jarrett, 1991; Deák, Flom, & Pick, 2000; Moll & Tomasello, 2004). Between 10 and 12 months of age, infants learn to not follow the head direction of people whose eyes are closed or covered (Meltzoff & Brooks, 2007). This might be indicating their rudimentary understanding of the function of eye gaze around this age, as they also start to engage in a protodeclarative pointing, which is thought to serve to direct other people's attention towards their own attentional target (Carpenter, Nagell, Tomasello, Butterworth, & Moore, 1998). This has been considered as a sign of infants starting to understand others as an intentional agent, and that their attentional target could be different from their own (Baron-Cohen, 1989; Brooks & Meltzoff, 2004, 2005; Meltzoff, 2007). It may be with such a recognition that infants start to engage in these gazing and pointing behaviours so other people's attentional target would shift and align with theirs. As such, infants can actively engage with and influence other people, rather than only being passive receptors of social information given from others.

Conventional methods of infant social cognition research: infants as receivers of information

The field of social cognition has aimed to understand and describe the cognition at play during social exchanges between individuals. The term 'social cognition' is loosely defined, and therefore, the field has addressed a wide range of questions. In the context of developmental social cognitive science, it typically concerns how infants can appropriately engage in complex exchanges of social cues in a social interaction in spite of their limited cognitive skills and social experiences (Reid & Striano, 2007).

Cognitive neuroscience has explored and shed light on the cognitive processes underlying infants' social attentional and communicative behaviour. The evidence from these studies has shown that, for instance, infants have separable brain regions that are engaged in initiating and responding to gaze exchanges (joint attention) (Mundy, 2018). Furthermore, research has shown that infants' ability to follow other people's gaze facilitates the encoding of other people's attentional target compared to other distractor objects (Hoehl, Wahl, & Pauen, 2014; Reid, Striano, Kaufman, & Johnson, 2004; Wahl, Michel, Pauen, & Hoehl, 2013). The evidence demonstrates that infants are equipped with neurobiological functions that enable them to take part in reciprocal exchanges of social cues.

Despite the field of social cognition primarily aiming to understand the working of cognition underlying social interactions where social partners influence each other, the experimental paradigms routinely used in cognitive neuroscience research on this topic do not typically encompass *reciprocal* interactions. Thus, we do not yet know what cognitive processes underlie bidirectional exchanges of social signals, and how infants cognitively manage their participation in a social interaction. Typically, studies using neural measures present infants with pictures or simple videos showing a person or virtual character looking towards certain objects near their own face (e.g., Grossmann & Johnson, 2010; Grossmann, Lloyd-Fox, & Johnson, 2013; Michel et al., 2015). In such a paradigm, the infants' role is only to respond to the presented stimuli, and therefore they would not affect the stimuli. As there would be no reciprocal interaction between the infant and the stimuli presented, evidence from such a paradigm is likely to only represent a part of what happens in a real-life experience. Hence, it is difficult to discuss the neural and cognitive mechanisms underlying a social interaction between infants and others based on these reports. An alternative line of research has addressed this question by interpreting correlations between infants' baseline neural activity and their social behavioural skills assessed in a live social interaction (e.g., Mundy, Card, & Fox, 2000). These studies are informative in demonstrating how typical neural activity patterns in infants are related to infants' social behaviour, and have shed lights on infants' cognitive abilities that underpin their social skills (Perone & Gartstein, 2019). Yet, the data of focus (i.e., baseline neural activity) in these studies are typically collected separately from the interaction phase where infants' social behaviour is evaluated. Hence, it is difficult to examine the *online* neural or cognitive processes underlying social interactions or infants' social behaviour. Therefore, even with these studies, less is known about the neural mechanisms of how infants actively engage in social interactions and contribute to the reciprocity of social interactions.

Despite these open questions, a large body of evidence from developmental cognitive neuroscience has certainly been informative, suggesting infants' predisposition towards social information and leading to the development of theories of infant social cognition (e.g., Grossmann, 2015; Reid & Striano, 2007). For instance, the directed attention model was proposed to address the gap in the existing models of social cognition which are primarily focusing on adults with more mature cognitive abilities and more social experiences (Reid & Striano, 2007). The model describes infant social cognition processes sequentially, from the detection of biological and/or socially-relevant agents to understanding the goal and intention of the agent's behaviour. Similarly, the social brain framework proposed by Grossman (2015) describes infants' perceptual sensitivity and predispositions towards social stimuli, and how the integration of multimodal perceptual cues could lead to the understanding of others. As is the case with the directed attention model, his account discusses how behaviouristic, perceptual understanding leads on to cognitivistic semantic understanding (Reddy, 2008).

Whilst these theories provide an insightful and plausible picture of infant social cognition, considering the developmental trajectory of perceptual and cognitive abilities, they do not answer the fundamental issue of what neural and cognitive mechanisms underpin the dynamic exchanges of social cue. This is largely because, as discussed above, the evidence on which these accounts are based almost predominantly come from studies with an experimental setting involving no reciprocal social interaction. In these studies, infants' positions are almost predominantly passive receptors of social information. Yet, the existing evidence shows that infants can engage with other people through various social attentional behaviour such as protodeclarative pointing (Baron-Cohen, 1989; Carpenter et al., 1998). Therefore, the models require expansion so that they could incorporate explanations of how infants learn to engage with other people in a social situation beyond the processing of social information. Furthermore, these theories fail to consider social environmental factors that inform other people's behaviour, whilst it has been reported that infants show differential behavioural responses to different contexts, even when the overt action of the social agent's is the same (Gergely, Bekkering, & Király, 2002). Hence, it is plausible that, whilst perceptual features of social agents are likely to play a role in infant cognition and attentional processes (Grossmann, 2015; Reid & Striano, 2007), infants are capable of processing and interpreting information about not only the observable behaviour of other people, but also the surrounding environment and situational factors. The lack of consideration of these contextual factors is a further consequence of the conventional experimental paradigm where infants are simply presented with stimuli in a controlled lab setting where little or no contextual factors to define the presented stimuli.

To sum up, the question of what neural and cognitive processes underlie infant-adult reciprocal social interactions remains open despite ample studies taking the cognitive neuroscience approach. This points to the constraints of the current methodology of cognitive neuroscience, where, in many cases, infant neural responses are not yet recorded whilst infants engage in a social interaction with others. This creates a gap between the methods and the aim of the research field of infant social cognition, as the key goal of the field includes to explore how infants engage with others in a social situation competently despite their limited cognitive skills and social experiences.

Disparity between the core questions and methodology of social cognition: 'third-person' 'isolation' approach of modern science

Where does such a disparity come from, between the key question of the research field investigating infant social cognition, and its research methodology? To address this question, it is important to look to the theoretical framework underpinning the methodologies of modern science and the key assumptions which inform them. One important framework is reductionism, which states that a complex system, including human cognition, can be explained as a set of isolatable components (Jessor, 1958; Richardson, 1979; Sawyer, 2002). Each component should be observable and measurable, and hence, it is possible to understand the complex system in a mechanistic way (Hull, 1943). In the context of social cognition research, it is typically assumed that the unobservable psychological elements of humans (e.g., thoughts, feeling, knowledge) need to be inferred from the observable physical components of humans (e.g., behaviour) (Asch, 1952; Costall, 2006; Costall, Leudar, & Reddy, 2006; Neisser, 1980). According to this reductionistic and mechanistic understanding of human cognition, social situations, human cognition, and behaviour are individual components, each of which acts independently of other components. Cognition can thus be studied on its own via inferences from behaviour, regardless of the situation where they emerge. More specifically, 'social' cognition can be studied even when participants are not engaging in a social situation, as long as they are presented with social stimuli. As such, under these assumptions, social environmental and contextual factors that define social stimuli are often not discussed. These are reflected in a typical experimental setup of cognitive neuroscience, in which participants are often asked to observe social stimuli (e.g., faces) and think about the agent's mental states or experiences, rather than take part in a social interaction with other people (Gallotti & Frith, 2013; Schilbach et al., 2013).

Another fundamental assumption of modern scientific method is that understanding can be gained through observation of the phenomenon (Becchio, Sartori, & Castiello, 2010). Therefore, a scientist acts as an observer, who does not engage in the phenomena itself, and the ‘spectatorial knowledge’ that can be gained via observation is the knowledge to pursue (Asch, 1952; Neisser, 1980). The idea that knowledge can be gained from an observer point of view has been incorporated in the field of social cognition in such a way that we understand other people’s minds as an observer, and how we think about others as an observer is the focus of research. According to this idea, cognition is typically understood as a one-way process in an individualistic fashion, in such a way that ‘a detached observer reads the mental states of another person, who, in turn, is not affected by this and cannot react to it’ (Schilbach et al., 2013, p.396).

These assumptions have informed the ‘third-person’ paradigm, or ‘isolation’ paradigm, in which the participant’s position is predominantly an observer outside of the social context (Schilbach et al., 2013). The key characteristics of the ‘third-person’ approach across research investigating social cognition are (1) the isolation from the social environment or interaction, and (2) the assumption that human cognition works as an independent component and one’s cognition does not influence or is not influenced by another person’s cognition or environmental elements. When the focus of research is to examine how social information is processed when it is presented, these paradigms are effective and can provide insightful knowledge. Yet, when the question concerns the dynamic cognition that supports active interactions with other people and social environment, these methodologies need to be different. The disassociation between the key question and the methodological approach is also present in the research field of infant social cognition. Hence, a different paradigm is needed so that researchers can explore the dynamics of cognition that underpin and develop through a reciprocal interaction between an infant and an adult (Hoehl & Markova, 2018). The paradigm requires a shift towards a more naturalistic setting rather than a conventional controlled lab setting, where infants engage in a reciprocal interaction with other people rather than being presented with pre-recorded stimuli.

Call for a paradigm shift: the ‘second-person’ approach of cognitive neuroscience

The ’second-person’ approach of cognitive neuroscience has been proposed with a motivation to resolve the conflict between conceptualisation of social cognition and methodological stances of the research field of cognitive neuroscience investigating social cognition (Schilbach et al., 2013). The central thesis of the second-person approach is that the brain activity needs to be examined in a paradigm where participants take part in a dynamically unfolding social interaction (Schilbach et al., 2013).

The framework underlying the second-person approach is largely related to the embodied cognition approach (Gibson, 1986; Thompson, 2007). The embodied cognition approach views perceptual information processing as an active process executed by an organism situated in the environment (Thompson, 2007; Varela, 1997; Weber & Varela, 2002). Participants are embedded in and interact with the perceived world, as cognition embedded in the bodily organism is an active interaction between perception and the organism's action potential (Gibson, 1986; Thompson, 2007). In this sense, the environment, behaviour and cognition are not separate entities, and each works in relation to another, rather than independently.

The second-person approach also questions the individualistic view that cognition is a one-way process, where one's cognition is detached from other people's cognition (Schilbach et al., 2013). Instead, it proposes the enactivist position to argue for a relational and dynamic account of cognition (Gallotti & Frith, 2013; Siposova & Carpenter, 2019). That is, social interactions are not simply the output of cognitive processes taking place independently in each individual, as cognition of each person is regulated by the dynamic encounters with the surrounding physical and social environment including each other's cognition (De Jaegher & Di Paolo, 2007). This creates an joint and interactive unit that cannot be reduced to the attributes of individual minds (De Jaegher & Di Paolo, 2007; Gallotti & Frith, 2013; Siposova & Carpenter, 2019). Interactionists holds a similar theoretical standpoint that representation, including other people's mind, emerges in dynamic anticipatory processes that only unfolds within a moment-by-moment social interaction (Bickhard, 2009; Michael, 2011). When taking these theoretical approaches, human cognition develops *within* social interactions, in which social partners influence each other creating a unit. Therefore, human cognition needs to be studied while participants actively engage in a reciprocal social interaction with other people, rather than observing others from a position detached from the social context.

Before continuing on, it is important to acknowledge that there are different theoretical backgrounds that coincide with the second-person approach, including enactivism (e.g., De Jaegher & Di Paolo, 2007) and interactionism (e.g., Bickhard, 2009; Michael, 2011). Whilst they differ from one another in various ways, the detailed discussion of each claim is beyond the scope of this introductory section of this thesis (for discussions, see e.g., Bickhard, 2016; De Jaegher & Di Paolo, 2013). Nonetheless, despite such variations, the consensus is that human social cognition and the neural mechanisms that support human social interactions need to be studied in a dynamic paradigm where participants engage with others. In this thesis, the term '*second-person' approach* of cognitive neuroscience generally refers to the overarching standpoint that puts forward a proposal of studying human social cognition in a paradigm where research participants engage in a reciprocal social interaction with other people rather than simply observe other people from outside of the interaction.

Infant cognition from the second-person perspective: how is it different from infant cognition from the third-person perspective?

The importance of taking the second-person approach is not only due to the definition of social cognition being contextual and relational, but also because of the evidence that suggests behavioural differences between the second-person and the third-person perspectives. It has been proposed that when one is engaging with another person in a social situation, acting as the second person, they have access to more situational and subtle information that indicates the social partner's mental states, and also become more emotionally involved, as compared to when they are observing others without interaction (Butterfill, 2013; Schilbach et al., 2013; Siposova & Carpenter, 2019). These lead to a broader understanding of the social partner's behaviour and of options available for action (Gallotti & Frith, 2013; Siposova & Carpenter, 2019). These proposals are generally made to discuss adults' cognition in a social interactional situation, who have more mature skills to understand other people's minds. The same question could be asked about infants: is the working of infant cognition different depending on their position relative to others in a social interactional situation?

The evidence from behavioural studies suggest that infants do behave differently depending on whether they are actively engaging in an interaction or acting as an observer of an interaction between others (e.g., Moll, Carpenter, & Tomasello, 2007; Moll & Tomasello, 2007). Differential behaviour might indicate differential workings of cognition depending on the infant's position in relation to other people in each situation. This means that what we know about social cognition from the third-person paradigm might need refining when we consider infant cognition at play in a social interactional setting. Infants' differential behaviour in different social positions has been reported by Moll and colleagues (2007). They investigated whether infants' understanding about other people's visual experience (i.e., what they have seen and have not) is affected by whether or not they have been in direct interaction. They compared two different situations. In the first, infants directly interacted with an experimenter playing with two toys, and in the second, they simply observed the experimenter playing with two toys with another person (Moll et al., 2007). In both situations, the experimenter then left the room, and infants were shown a third toy. The experimenter then returned, expressed interest and excitement in the toys, and asked infants to pass a toy without specifying which toy they wanted (e.g., no gaze cues or pointing). In the situation where infants had directly interacted with the experimenter, they passed the novel toy, indicating their understanding that this toy, being novel, would be of greater interest to the experimenter. In the situation where infants had only observed the experimenter's prior interactions, the selection of toys to pass was more random, suggesting that they did not understand which one was novel or more interesting to the experimenter (Moll et al., 2007). This study highlights that direct social interaction makes a critical difference in the working of infant social cognition and their understanding of other people's mental experiences. Prior to this study, it had been proposed that infants of 14 months of age need to observe the target person actively engaging with a target object in order to be able to infer whether the person had seen it in the past (Tomasello & Haberl, 2003; Woodward, 1998, 2003). Yet, the results of Moll et al. (2007) have suggested an alternative account that infants are able to infer what another person has witnessed as long as they have been actively engaging with the infant, even when they had not had any physical active engagement with the object (e.g., touching, manipulating). If infants were not positioned as the second-person to the experimenter, such inferences would be challenging even for 18-month-olds (Tomasello & Haberl, 2003). Taken together, being in the second-position person gives us, adults and infants alike, unique access to information that cannot be obtained from the observer perspective (Butterfill, 2013; Gallotti & Frith, 2013).

In a similar vein, another line of research has investigated whether infants' proximity to other people influences their sensitivity to social information. These studies have shown that infants are only sensitive to social communicative cues when they are in an interactional context. As mentioned, one of the most important social cues for infants is eye gaze. Receiving direct eye contact indicates that the receiver is being addressed as “you” (i.e., the second person) (Kampe, Frith, & Frith, 2003). Being addressed as the second person enhances an infant’s affective engagement with the social partner, as infants as young as 3 months of age smile in response to eye contact and decrease smiling when a partner's gaze is averted (Hains & Muir, 1996). Being addressed as the second person is also known to modulate infants’ subsequent attentional engagement. Farroni and colleagues (2003) conducted a series of studies where 4-month-old infants were presented with an adult face whose gaze is directed to an object placed next to the face. It was reported that infants only followed the gaze direction when they saw the adult gaze being directed at them (Farroni et al., 2003). Senju and Csibra (2008) reported a similar effect with 6-month-old infants. Given such a high sensitivity towards eyes that infants demonstrate, Beier and Spelke (2012) investigated whether infants’ sensitivity towards other people’s gaze is present when infants are positioned as a third person. The result suggested that 9-month-old infants are not able to differentiate direct and averted gaze between two individuals from the observer perspective, or follow the gaze of another person participating in an interaction with someone else. These results indicate that young infants are only sensitive to the gaze direction of others and its function when they are positioned as a social partner of the other person, but not when they are outside of direct interaction. Similar effects of being in the second-person position in a social situation has been reported in the context of learning novel words (Floor & Akhtar, 2006) or actions (Herold & Akhtar, 2008). These studies consistently demonstrate what they are sensitive to in a social interactional context might not always inform their cognition and behaviour when they are in an observer position outside of a social interaction. This again highlights the importance of studying infant social cognition in an interactional setting where infants can actively engage in a social interaction with others, rather than the conventional observational setting.

To sum up, the second-person approach, where infant cognition is studied while they are actively interacting with others rather than being an observer of others, has a theoretical rationale and empirical support. Conceptually, infant cognition and its development are intertwined with a quality of social interaction that infants experience. This raises a question about the use of conventional third-person paradigms of cognitive neuroscience research exploring infant social cognition outside of social interactional settings, as it might only capture a specific aspect of infant cognition, overlooking the whole picture. Empirical evidence suggests that infant cognition works differently depending on their position and perspective in relation to other people (second-person or third-person), as reflected in their differential behaviour across different interactional contexts.

Methods of second-person cognitive neuroscience for developmental cognitive psychology

How could we study infant social cognition taking the second-person cognitive neuroscience approach? There has been a surge in the last few decades in the number of studies that investigated adult social cognition taking this approach. These studies have provided evidence to support the concept that neural processes are different when participants are positioned within a social interaction taking the second-person perspective as compared to when they act as an observer taking a third-person view (Redcay & Schilbach, 2019). It is hence promising and important to conduct the same line of cognitive neuroscience research with infant participants to shed light on neural and cognitive processes that uniquely emerge in a social interactional setting, and how these processes develop from infancy to adulthood.

The relative lack of the second-person cognitive neuroscience research investigating infant social cognition is largely due to methodological constraints related to the proneness of brain imaging methods to gross motor artefact. We cannot ask infants to remain still during the data acquisition, and therefore many studies have used a paradigm where infants only need to attend to the screen in front of them, encouraging minimal movement. In transitioning towards a more naturalistic experimental setting, the functional magnetic resonance imaging (fMRI) is the most challenging method to use with infant participants, due to the motion restriction and a high level of noise that occurs during fMRI scanning. Electroencephalogram (EEG) and functional near-infrared spectroscopy (fNIRS) have become more common methods for infants as compared to fMRI (Nguyen, Bánki, Markova, & Hoehl, 2020; Wass, Whitehorn, Marriott Haresign, Phillips, & Leong, 2020). Whilst fNIRS is known to be relatively more robust to gross motor movement than EEG, EEG has been used widely with infant participants and provided insights into infants' cognitive processing of social situations (de Haan, 2013; Hoehl & Wahl, 2012). One of the features of EEG that is important and relevant to the second-person approach is its fine temporal resolution (Cohen, 2011; Czeszumski et al., 2020). The temporal resolution of EEG is better than that of other brain scanning methods such as fMRI and fNIRS, as EEG measures changes in electrical activity that occur more rapidly than metabolic changes in blood oxygenation levels, which are what is captured by fMRI and fNIRS (Cohen, 2011; Logothetis, 2008; Pinti et al., 2018). Behaviour in social interactions has smooth, contingent rhythmic patterns, which are co-created by social partners as they adjust their behaviour according to others (Murray & Trevarthen, 1985; Reddy, 2008; Stern, 1985). Therefore, the neural mechanisms that underpin such behavioural coordination should also be represented by rhythmic activity, and the measure we use should be able to capture the temporality with fine resolution. Furthermore, prior research has provided evidence that certain EEG components, for instance, event-related potentials (ERPs) and event-related oscillations (EROs), have been linked to certain cognitive and perceptual processes (Azhari et al., 2020). Such knowledge has enabled the identification of the neural correlates of infant cognition. Whilst the susceptibility to movement is a major disadvantage of EEG, the recent advances in removing motor artefact have enhanced the usability of EEG in a less controlled setting as well as the development of portable devices (Georgieva et al., 2020; Noreika, Georgieva, Wass, & Leong, 2020). Thus, the application of EEG has a large potential to be utilised in the second-person cognitive neuroscience investigating infant social cognition.

Moving towards a live experimental setting: single-brain approach

In transitioning towards a less controlled social setting where infants can engage with others in a naturalistic manner, we must review and take forward what we have learned so far from conventional paradigms that have utilised screen-based stimuli but which also have begun to take steps towards more naturalistic methods with the use of live presentation. Whilst studies with adult participants have utilised screen-based methods to create a reciprocal and contingent exchange between two participants (e.g., Saito et al., 2010), it is important for infant studies to employ a live interaction. This is because previous research has suggested that infants process screen-based and live presentation, or 2D and 3D images presentation differently, as they show differential behavioural responses according to the presentation modality (Barr, Muentener, Garcia, Fujimoto, & Chávez, 2007; Cleveland & Striano, 2008; Shimada & Hiraki, 2006).

A few studies have implemented a live joint attentional situation in an experimental setting (Hoehl, Michel, Reid, Parise, & Striano, 2014; Parise, Reid, Stets, & Striano, 2008; Striano, Reid, & Hoehl, 2006). Infants in these studies were in the second-person position relative to the experimenter, as they were engaged with by the experimenter in a live manner. For instance, to investigate the effect of joint attention on object processing, Striano and colleagues (2006) as well as Hoehl and colleagues (2014) compared infants' neural responses to an object presented on a screen in a condition where an experimenter engaged in joint attention with infants in a live format, and a condition where an experimenter only attended to an object shown on the screen. Both studies found the effect of joint attention on the processing of the object as reflected in the increased amplitude of the Negative Central ERP component (Nc) (Striano et al., 2006) and the desynchronisation in alpha power (Hoehl et al., 2014). Whilst these studies are among the first to use a live 'interactional' paradigm, the focus of the analysis in each study was infants' processing of the object projected on a screen. Therefore, infants' engagement in a social interaction with the experimenter was not the focus, and the online neural activity while infants engage in joint attentional episodes with the experimenter has not yet been explored. Critically, the experimenter's looking behaviour in these studies was highly controlled for the experimental purposes (i.e., they looked to the infant and the object in an alternate manner) and hence, infants are likely to have received few or no contingent responses. These settings lack in a reciprocal interaction between an infant and an experimenter, making it challenging to investigate the neural mechanisms underlying reciprocal social interactions from their data.

The studies referred to above have taken the 'single-brain' approach, whereby the data of interest are collected from one individual at a time (Redcay & Schilbach, 2019). Many aspects of social interaction could be studied with this approach (e.g., Stephens, Silbert, & Hasson, 2010). Accordingly, many existing studies taking a second-person approach with adult participants have used this single-brain approach. This is probably because of its similarity to the standard procedure of conventional cognitive neuroscience focusing on an observer's brain activity (Redcay & Schilbach, 2019). Much research in this category examined a neural response to a social stimulus, such as faces or eyes, typically during an interaction with a trained research assistant, or, especially in adult research, a virtual agent that can show contingent and spontaneous responses to participants (Noreika et al., 2020). Such experimental controls, seen in a highly structured interaction, are a common feature of both the single-brain second-person approach and the conventional third-person approach. Such commonality makes evidence from the single-brain second-person research comparable to knowledge from the conventional third-person research. This allows us to examine whether there is a systematic difference in brain activity between when we are interacting with a social partner and when we are simply observing them. This is an advantage of this single-brain second-person approach, and it can provide insights into the core social interactive processes that cannot be probed outside of a social interactive context.

Moving towards an assessment of interacting brains: simultaneous dual-brain approach (hyperscanning)

Increasing focus on the second-person approach of social cognition research in the field of cognitive neuroscience asks what neural mechanisms underlie social interaction (Schilbach et al., 2013). During a social interaction, social partners adjust the timing and intensity of their communicative behaviour according to the other's actions, and such a mutual coordination of behaviour is considered a critical element of effective human communication (Csibra, 2010; Feldman, 2007; Fiske, 1993; Schilbach et al., 2013). Yet, little is known about whether, and how, such a mutual coordination happens at a neural level (Hari & Kujala, 2009; Schilbach et al., 2013). In order to investigate whether such 'brain-to-brain' coordination underlies the 'behaviour-to-behaviour' coordination that emerges during a social interaction, it is critical to examine the neural activity of two individuals at the same time whilst they engage in a live *reciprocal* interaction (Babiloni & Astolfi, 2014b; Nguyen et al., 2020; Wass et al., 2020). This method has been termed 'hyperscanning' (Montague et al., 2002), or 'simultaneous dual-brain approach' (Redcay & Schilbach, 2019).

To date, research utilising hyperscanning methods has typically been conducted with adult participants (for scoping reviews, see e.g., Babiloni & Astolfi, 2014; Redcay and Schilbach, 2019), and there have only been a handful of studies that utilised a hyperscanning EEG technique with infant participants. This research has been led exclusively by two research groups so far (Leong et al., 2017; Santamaria et al., 2020; Wass, Noreika, et al., 2018). The types of interaction have included semi-structured interactions between an infant and an experimenter (Leong et al., 2017), as well as a play session between an infant and their parent (Santamaria et al., 2020; Wass et al., 2018). These studies are the first of their kind whereby infants were positioned as the second person and could experience a reciprocal interaction with a live social partner.

The hyperscanning EEG technique with infant participants has enabled the investigation of what social cues might lead to the alignment of neural activity between an infant and an adult. For instance, it has been investigated whether direct eye contact would facilitate the alignment of an infant's and an adult experimenter's neural activities (Leong et al., 2017). The infant's and the experimenter's neural activities were simultaneously collected using EEG whilst the experimenter sang a nursery rhyme to the infant. The results suggested that direct eye contact led to a higher degree of brain-to-brain coordination (i.e., temporal alignment of two neural activities across participants) compared to the control condition where there was no direct eye contact between an infant and an experimenter. Whilst the interaction that took place between an infant and an experimenter was rather restricted (i.e., during the interaction, the experimenter sang a song to them from a distance), it was reported that infants responded to the experimenter by vocalisation particularly when they had a direct eye contact as compared to when they did not (Leong et al., 2017). Notably, such a difference in infants' vocalisation across eye-contact conditions was not observed when infants were presented with a pre-recorded video of the same experimenter singing the same song. This indicates a differing degree to which infants attempted to interact with the experimenter depending on whether it was a live or a screen-based interaction. This points to the importance of simultaneously measuring brain activities of infants and adults during a live interaction to provide knowledge of high ecological validity regarding the brain mechanisms underlying adult-infant social interactions. Following this, Santamaria and colleagues (2020) showed that parental positive emotional expression during an infant-adult play session also induced a higher level of interpersonal brain-to-brain coordination. The brain activity data were collected from an infant and their parent whilst they were playing together with an object in a naturalistic and reciprocal manner, demonstrating the feasibility of this simultaneous scanning of multiple brains including infant participants during a less constrained dynamic interaction. These studies provided the first pieces of evidence that infant brain activity is coupled with an adult's brain activity through social cues. This suggests that there are indeed neural underpinnings to the behavioural patterns that align during social interactions and that these are well-developed during infancy.

Beyond the findings of brain-to-brain coupling between infants and adults, an important question arising from these studies relates to the function of such neural coupling. Due to the data contamination from motor movement (Santamaria et al., 2020) or adult speech production (Leong et al., 2017), the analyses in these existing dual-brain second-person cognitive neuroscience with infant participants have been restricted to the data from central two channels only. This has made it difficult to examine the potential neural sources of the observed effects. Therefore, little can be discussed in terms of the cognitive activity underlying the observed brain-to-brain coupling and its function in infant cognitive development.

To address the issue, it would be optimal to start with developing a paradigm where the identification of neural markers of social cognition and social behaviour is possible. Utilising the technique of analysing ERPs or oscillatory activities would be beneficial, as ERP components and different frequencies of EEG have been linked to specific cognitive activities underlying certain activities or behaviour (de Haan, 2013; de Haan & Gunnar, 2009). Furthermore, the search for an optimal paradigm should begin with examinations using relatively controlled setting before moving onto a naturalistic setting, as the data might be more likely to be contaminated by artefact in a naturalistic and less controlled situation, and involve behavioural variations that could explain varying neural activities (Markova, Nguyen, & Hoehl, 2019; Nguyen et al., 2020). Indeed, Smith and colleagues (2021) investigated 6-month-old infants' neural oscillatory activity during a free play session with their mother with minimal constraints. Whilst the study identified different oscillatory activity patterns between social and non-social phases during a mother-infant free play session, the authors raised methodological and analytical concerns of using unconstrained interactional paradigms for infant EEG research (Smith et al., 2021). This included a large behavioural variation across dyads and a high level of noise in the data due to infants' movement. A large variability in behaviour could lead to difficulty in connecting electrophysiological characteristics to certain behaviour, resulting in the lack of interpretable results (Markova et al., 2019; Nguyen et al., 2020). Hence, we need to find a situation that can be reproduced in a lab setting but participants' behaviour is naturally less varied, such as a shared object play, because a certain level of structure would allow us to systematically code their behaviour, enabling the reliable identification of the relationship between neural markers and behaviour (Neale et al., 2018).

Addressing the question of the function of neural synchrony, a study by Wass and colleagues (2018) examined the relationship between visual attentional behaviour and theta-band neutral activity across infant-parent dyads. EEG data from infants and their parents were simultaneously collected during an object play session, and their looking behaviour were also monitored. The results showed that the parent’s theta power activity increased when infants looked toward the object, and that the infant’s attention was sustained on the object longer when their parent’s theta power increased in response to the infant’s looking behaviour. By examining both neural activities and looking behaviour as well as the link between the two, the study highlighted how infants and parents influence each other during a reciprocal and dynamic exchange at neural and behavioural levels. This study constitutes an important step towards the direction of more dynamic and naturalistic interactional paradigms that allow for an examination of infant cognition and brain mechanisms underlying a reciprocal social interaction. Based on the paradigm presented in this study, it is possible to explore the dynamic neural-behavioural processes that emerge during a social interaction. For instance, an open question arising from this study is regarding how parental neural sensitivity affects infant cognition and learning. This study demonstrated the attentional scaffolding effect of parental neural sensitivity, in such a way that parents’ sensitive monitoring and adjustment to infants’ attentional behaviour encouraged infant sustained attention. Attention is considered to be important for learning (e.g., Choudhury & Gorman, 2000), so we can assume that parental neural sensitivity might have some effect on infant learning in a social situation by scaffolding their attention. Future studies can investigate the effect of parental neural sensitivity on infant learning. Critically, such a specific hypothesis could be formed because this study by Wass et al. (2018) has associated specific behaviour (e.g., infant looking) with a certain neural activity. This provides a model for future studies to follow in order to advance our understanding about the neural dynamic processes underlying a reciprocal social interaction between an adult and infant.

Whilst the contribution of the study by Wass et al. (2018) to the field is certainly critical, the importance of establishing a stable platform which we can turn to still persists, so that we can better understand and appropriately interpret the data from a dynamic and naturalistic interaction. As was the case with other studies using the second-person dual-brain methods (Leong et al., 2017, Santamaria et al., 2020), this study also only examined two central electrodes to exclude data contaminated from gross movement. This limits the inferences that could be made about the cognitive processes underlying a social interaction. Nonetheless, this has provided groundwork in advancing our methodologies in moving to a dynamic interactional research setting. This means that we can begin to think about how we can balance the second-person research perspectives and the experimental control required to make a reliable inference, so we can better understand the neural mechanisms underlying dynamic and reciprocal social interactions between an infant and an adult.

Thesis objectives

The challenges identified here to conduct second-person research investigating infant social cognition indicate the critical importance of first establishing a platform that we can turn to when moving towards more naturalistic interactional settings, and ultimately carrying out research using a simultaneous dual-brain approach to address one of the main questions in the field of infant social cognition regarding the brain mechanisms which underlie a reciprocal infant-adult interaction.

This thesis has two primary objectives. The first is to establish and validate a protocol for data acquisition and analyses for research using EEG in a more naturalistic and dynamic setting, where infants can actively interact with another person in a live manner. The second is to explore the interplay between infant cognition and the social world, by implementing a less constrained and more dynamic paradigm. In doing so, the studies presented in the thesis jointly examine the necessary steps to take to move towards research taking the second-person cognitive neuroscience approach which advances our understanding of infant social cognition, whilst balancing the interpretability of the data and the naturalistic social nature of research paradigms.

Thesis structure

The thesis commences with a meta-analytic literature review (Chapter 2), examining the variation of methodology reported in published studies that have taken a simultaneous dual-brain second-person neuroscience approach, utilising hyperscanning methods. The results of this study highlighted the challenges that the field is facing: namely, a large heterogeneity in reported effect sizes found across published studies. This suggested that it is now timely to approach improvement in the comparability of research, such as establishing standardised methods and promoting open science practices including data and code sharing to achieve higher reproducibility. This would enable a more reliable interpretation of the reported results from hyperscanning research so we can better understand the cognitive function of brain-to-brain coordination that emerges during a social interaction. Furthermore, this meta-analysis cautions that we must ensure the quality of the data is adequate to conduct an interpretable and reliable analysis. Due to the dynamic nature of social environment, data collected from a paradigm which allows participants to interact together could suffer from a high level of noise and artefact in the data. This is particularly relevant when working with infants, who cannot stay still for a period of time or follow specific experimental procedures, and show high level of inter-subject and inter-trial variability (de Haan, 2013). This means that they only contribute a small number of data, and the number of usable data will be even smaller as contaminated datasets need to be excluded, which will affect the signal-to-noise ratio of the data (i.e., data quality). This indicates that we should first identify a reliable neural marker of social cognition and behaviour using a relatively controlled experimental setting before moving on to a less controlled dynamic paradigm, which is likely to yield noisier data that are difficult to interpret.

Therefore, from Chapters 3 to 6, four experimental studies will be presented where we examined infant social cognition in a manner that transitioned from screen-based to more naturalistic paradigms. Chapters 3 and 4 presented a screen-based paradigm where we aimed to reliably identify neural markers of infant cognition that occurs in a social environment. In Chapter 3, we explored how infants encode information differently from two adults who give gaze cues to a target object with different levels of accuracy. This study utilised a conventional event-related potential paradigm using screen-based static stimuli and identified the neural markers (ERP components) of infants' selective attention and cognition that emerge in a social situation. Chapter 4 presents a study which utilised a screen-based, yet more dynamic video stimuli and explored the neural processing of unexpected events. The study identified some challenges of using dynamic perceptual inputs as stimuli in EEG studies.

In Chapter 5, we transitioned into more naturalistic social contexts, and analysed infant cognition whilst infants were faced with an adult demonstrating novel object labels in a live interaction, utilising EEG oscillatory analysis methods. The analysis was conducted based on previous studies using a more controlled setting which reported neural correlates of cognitive processes of interest. The study not only showed the feasibility of research taking the second-person cognitive neuroscience approach with infant participants, but also advanced our knowledge about infant word learning a step further, by demonstrating the trajectory from the encoding of lexical semantic information to its consolidation as knowledge. Similarly, in Chapter 6, we explored infants' cognition by analysing neural data collected from infants who engage in a traditional Piagetian A-not-B search task (1954) in a live format, with the aim of identifying systematic differences in neural activity between infants who make perseverative reaching errors and those who do not. This study was, to our knowledge, the first of its kind to validate the feasibility of utilising neurophysiological measures in this traditional interactive behavioural paradigm, in such a way that it does not interfere with its standard procedure.

Despite some challenges, these studies jointly demonstrate the potential for conducting research in a more dynamic setting that investigates infant social cognition taking the second-person cognitive neuroscience approach to advance our knowledge about the intricate interplay between infant cognition, behaviour, and the social environment. Due to the global pandemic that occurred during the time of this research project, it was not possible to present a series of studies in a sequential manner that lead to a hyperscanning study where an infant-adult dyad’s brain activities are simultaneously monitored during a naturalistic social interaction, whilst this is thus far considered as an optimal environment to study second-person cognitive processes. Nevertheless, the thesis presents an exemplary pathway towards a dynamic social interactive paradigm taking steps from a relatively controlled paradigm so that the data can be reliably interpreted in terms of social cognition and behaviour that uniquely occur during a reciprocal and dynamic social interaction. The thesis concludes by addressing some challenges of such an approach, to which solutions are proposed, as well as discussing future directions for the field.

Chapter 2

A systematic meta-analytic review of methodological factors in hyperscanning studies in social contexts

Chapter Introduction

To better understand the neural mechanisms underlying reciprocal infant-adult social interactions, the use of hyperscanning techniques is optimal, as it can examine the dynamic relationship between brain activities of two social partners. In infancy research though, there have been few studies investigating brain-to-brain coordination that emerges between infant-adult dyads during social interactions utilising this technique with electroencephalogram (EEG). Yet, the existing studies (e.g., Leong et al., 2017, Santamaria et al., 2020) had to focus their analyses on the data only from two central electrodes due to the motor artefacts. Therefore, the source of the observed cross-brain coordination could not be reliably identified, making it challenging to discuss the cognitive function of such inter-brain coupling.

To consider more optimal study and analysis designs, it is helpful to systematically review the literature including studies with adult participants. In this chapter, we present a meta-analysis reviewing published studies that utilised hyperscanning techniques to investigate brain-to-brain coordination that emerges between interacting individuals during social interactions.

Abstract

The introduction of hyperscanning techniques, concurrent measurement of brain activities of two or more individuals engaged in interactive contexts (Montague et al., 2002), has brought us closer to investigating the neural networks that support human dynamic social interactions. The last few decades have seen a surge of published papers in this domain, and although challenging, hyperscanning is shown to be feasible. Many studies have reported synchronised patterns between two brain activities ('inter-brain coupling'), and claim that these patterns specifically emerge *due to* social interactions. Yet, there have been concerns about whether the inter-brain coupling reported is really due to the social context, as it might be, at least in part, a methodological artefact (Burgess, 2013; Hamilton, 2021), or related to participant characteristics (Baker et al., 2016; Cheng, Li, & Hu, 2015; Li et al., 2020). To assess the influence of variations in study designs and analytic procedures on the reported degrees of inter-brain coupling, we reviewed the methodological and analytic parameters reported in the published studies utilising hyperscanning techniques. The three-level meta-analysis found no significant methodological factors that explained the variability of effect sizes reported, in terms of participants, experimental tasks or analytic procedures. As the variance of reported effect sizes were found to be attributed to the difference across studies rather than across different methodological choices, it is possible that the heterogeneity in the reported findings might be associated with researchers' degrees of freedom in choosing research methods, including different procedures of preprocessing data. The potential ways to address the issues are discussed, focusing on how we could move towards the standardisation of protocols, as well as ensuring transparent research practice.

Introduction

Social interactions play a critical role in human society. It is prevalent in our daily life, taking a variety of forms including, but not limited to, conversation, cooperation, competition, imitation, as well as teaching and learning. Therefore, it is no surprise that our behaviour, mental state, and brain activity are influenced by the physical presence and behaviour of others, as well as our understanding of other peoples' minds (Hari & Kujala, 2009). For the last few decades, a growing amount of research has been conducted based on the concept of 'second-person' neuroscience, which assumes that social cognition during a social interaction should be fundamentally different from social cognition emerging in a non-interactional context (Schilbach et al., 2013). In the second-person perspective, it is critical that neural data should be collected simultaneously from individuals who are engaging in a social interaction, as their neural activity might mutually influence each other during reciprocal interactions (Redcay & Schilbach, 2019).

The introduction of hyperscanning techniques has enabled us to investigate such 'interpersonal neural interactions' occurring during a social interaction (Montague et al., 2002). The term "hyperscanning" was originally proposed as a method to obtain two functional magnetic resonance imaging (fMRI) datasets from two individuals concurrently whilst they are interacting (Montague et al., 2002). Following this, the term has also been used in studies using other brain imaging techniques such as electroencephalography (EEG) and functional near-infrared spectroscopy (fNIRS). The term is currently used to generally describe a situation where brain imaging data are simultaneously collected from two or more individuals. Whilst many aspects of social interaction could be studied in experimental settings where the brain activity of just one individual is monitored (e.g., Stephens et al., 2010), the hyperscanning technique is considered to be optimal to investigate the alignment of brain activities in two individuals in terms of frequency and time (Babiloni & Astolfi, 2014; Nguyen et al., 2020; Wass et al., 2020). The last few decades have seen a surge in the number of papers published using this hyperscanning technique, which demonstrates that, although challenging, hyperscanning is feasible across different experimental paradigms, as well as across different participant groups including children and infants (Babiloni & Astolfi, 2014; Nguyen et al., 2020; Wass et al., 2020).

An fMRI has been one of the commonly used methods to explore the relationship between two brain activities arising during a social interactional context. A fMRI measures brain activity by recording changes associated with blood flow (blood-oxygen-level-dependent [BOLD] contrast), and can demonstrate the entire network of brain areas engaged while participants undergo specific tasks (Logothetis, 2008). One of the first studies using hyperscanning fMRI was reported by King-Casas and colleagues (2005). They connected scanners in California and Texas via the Internet to examine two participants' brain activities during an economic 'trust' game. One of the participants ('Investor') was given some 'funds', and asked to decide how much to invest in their partner ('Trustee'). The authors not only identified the specific brain regions engaged during the game, but also found that as the Investor learned to trust their partner, the relationship between the Investor's and the Trustee's brain activities shifted, in such a way that the activation patterns of the two brains became aligned within a shorter time frame (King-Casas et al., 2005). This study demonstrated that interacting individuals influence each other at a neural level, and such a brain-to-brain interaction develops according to the context of social exchanges. Following this, a body of research has investigated the neural correlates of social interactional behaviour, such as eye contact and joint attention (e.g., Koike, Sumiya, Nakagawa, Okazaki, & Sadato, 2019; Koike, Tanabe, et al., 2019; Saito et al., 2010) as well as speech production and comprehension (e.g., Spiegelhalder et al., 2014). Yet, a critical caveat of fMRI research remains in that participants are required to stay still in a laying position in a scanner. Furthermore, participants need a monitor to interact with their partner in a different scanner, meaning a face-to-face interaction cannot be incorporated in hyperscanning fMRI. Thus, the findings from fMRI research have limited ecological validity in their application to the study of social interactions given their face-to-face nature.

An EEG has been used more commonly as a method of hyperscanning in a more naturalistic setting than an fMRI, as well as traditional screen-based paradigms, due to its wide availability, a relatively low cost, and portability (Burgess, 2013; Czeszumski et al., 2020). An EEG measures changes in electrical activity in the brain that occur more rapidly than metabolic changes in blood oxygenation levels, which is what an fMRI captures (Nunez & Srinivasan, 2006). Therefore, it has a better temporal resolution than an fMRI. This is another advantage when studying social interactions that evolve on a fast scale (Cohen, 2011; Czeszumski et al., 2020). Dumas and colleagues (2010) were among the first to investigate the relationship between the brain activities of two individuals, using hyperscanning EEG techniques. In this study, participants were paired, and one of them ("Imitator") was instructed to imitate their partner's ("Model") spontaneous hand movement. The Model's hand movement was video-recorded and live-presented to the Imitator via a monitor. The authors found that the brain activities of the Imitator and the Model were synchronised, suggesting the coordination between neural activities occur during behavioural coordination (Dumas, Nadel, Soussignan, Martinerie, & Garnero, 2010). Similar to this study by Dumas et al. (2010), a number of hyperscanning EEG studies use paradigms where participants are asked to engage in a monitor-based task together rather than a face-to-face interaction (e.g., Astolfi et al., 2020; Babiloni, Astolfi, et al., 2007; Babiloni, Tocci, et al., 2007). The use of a screen limits the ecological validity of the experimental situation. Addressing the issue, a growing number of studies have been conducted with a more naturalistic paradigm where participants can engage in a face-to-face interaction. For instance, Lindenberger and colleagues (2009) have investigated synchronised neural activities across eight pairs of guitarists playing a short melody together (Lindenberger et al., 2009). The study showed that interpersonally coordinated actions (i.e., playing the short melody together) were preceded and accompanied by synchronised brain activity between a dyad. Furthermore, there have been studies which investigated the neural synchronisation between more than two brains (Astolfi et al., 2010; Dikker et al., 2017), and even outside of conventional laboratory settings such as school classroom (Dikker et al., 2017). Examining the behavioural correlates of the observed inter-brain synchronisation, these studies have proposed the function of inter-brain synchronisation should be linked to behavioural and attentional coordination (Dikker et al., 2017; Lindenberger et al., 2009).

An fNIRS is an alternative technique that has been used for hyperscanning research. Similar to an fMRI, an fNIRS measures blood oxygenation as an indirect index of brain activity (Pinti et al., 2018). Yet, the data recording is restricted to anatomically superficial brain areas, limiting its spatial resolution as compared to that of fMRI. Furthermore, its temporal resolution is not comparable to that of EEG. Despite these shortcomings, its high portability and resistance to motion artefacts have made fNIRS the most commonly-used measure in hyperscanning research to date (Hamilton, 2021; Pinti et al., 2018). One of the first studies using hyperscanning fNIRS was conducted by Cui and colleagues (20112), who analysed fNIRS data collected from pairs of participants who engaged in computer-based cooperation and competition tasks (Cui, Bryant, & Reiss, 2012). On a cooperation task, the participants were asked to press the button at the same time responding to a signal displayed on a shared monitor, whilst during a competition task, they were asked to press the button faster than the other participant. As a control, participants also completed the same task in an individual manner, where one participant was instructed to press the button as fast as they could while their partner observed. The authors found a greater degree of neural alignment between paired participants during cooperation than competition or individual performance (Cui et al., 2012). Following this study, a body of research using fNIRS hyperscanning has reported a synchronised activity between two brains during various tasks including verbal communication (e.g., Jiang et al., 2012; Nozawa, Sasaki, Sakaki, Yokoyama, & Kawashima, 2016) and motor coordination tasks (Holper, Scholkmann, & Wolf, 2012; Scholkmann, Holper, Wolf, & Wolf, 2013). These findings highlight in what context such brain-to-brain synchronisation increases compared to other contexts.

These findings of “brain-to-brain synchronisation” have led to claims that such a coupled activity pattern between two brains might be a “mechanisms for transmitting information” (Hasson, Ghazanfar, Galantucci, Garrod, & Keysers, 2012), or “mechanism of shared intentionality” (Fishburn et al., 2018). It has also been proposed that the coupled neural activity “might trigger the neural mechanism guiding social alignment” (Gvirts & Perlmutter, 2020). Yet, it is important to note that the existing reports have primarily focused on *when*, or *in what context*, inter-brain synchronisation occurs rather than *how* synchronisation occurs. Hence, much less is known about the function and processes of such brain-to-brain synchronisation. The primary question in cognitive neuroscience lies in what such a synchronised brain activity across individuals might tell us about human psychology that emerges uniquely during social interactions. Therefore, the question needs addressing further regarding how and why synchronisation emerges between two brains.

Before continuing, it is important to note about the terminology to describe the relationship between two brain activities. The alignment of multiple brain activities has been described with various terms, including “interpersonal brain synchronisation”, “inter-brain connectivity”, or “cross-brain coupling”. These terms appear to be used more as an umbrella term to explain various forms of association between multiple brains, rather than in a way that different terms refer to different types of inter-brain relationships. This is an issue that needs addressing as brain activity is typically captured as multiple dimensions, and accordingly a relationship between multiple datasets of brain activities can take different forms in different dimensions. For instance, synchronisation refers to two or more brains showing the same pattern of activity at the same time, whereas contingent alignment is also possible, where one brain activity following another brain's activity. These relationships can be evaluated by different types of measures. Nonetheless, in the current literature, differences in the terminology does not seem to always match differences in the types of relationships discussed. In this paper, from this section, we use 'inter-brain coupling' as a term broadly referring to multiple brain activities that are aligned in a certain manner.

Whilst many papers have reported inter-brain coupling that occurs during a social interaction, how to interpret such coupled neural activity between individuals is not straightforward. For this, it is important to review how inter-brain coupling has been measured, and what each measure of inter-brain coupling can capture. In research using hyperscanning fMRI, the relationship between two brain activities is typically examined using a correlation analysis (e.g., Saito et al., 2010; Shaw et al., 2018; Tanabe et al., 2012). A conventional correlation analysis typically focuses on a 'synchronisation' pattern between two datasets, with a larger correlation coefficient indicating a more similarity between the two. Critically, such a correlation analysis prevents the separate examination of temporal alignment in each phase (Shaw et al., 2018). This means that it is challenging to explore 'asymmetrical' patterns between two brain activities, which might be more prevalent as a pattern of social exchanges, in which different individuals can take different roles (Hamilton, 2021). For instance, the relationship between two time-series data (data containing multiple time points) can take a form of 'turn-taking', where two datasets share the same or similar dynamics but with a time lag. To evaluate such a relationship, a lagged correlation analysis has been used, in which one of the time-series data is moved forwards or backwards relative to the other dataset (e.g., King-Casas et al., 2005). Yet, either a correlation or lagged correlation analysis does not take into account how such a temporal relationship occurs. The measures themselves simply evaluate the similarity between two data. Hence, a higher correlation from such analyses can indicate coincidental synchrony. This makes it challenging to argue for a specific function of inter-brain coupling. With regards to hyperscanning EEG research, measures used to assess the degree of neural coupling between individuals vary across studies. Measures commonly used include covariance in amplitude or power, phase synchrony (Tass et al., 1998), or mostly Phase-Locking Value (PLV) (Lachaux, Rodriguez, Martinerie, & Varela, 1999), and Partial Directed Coherence (PDC) (Baccalá & Sameshima, 2001). Covariance measures a positive or negative association between two datasets as correlation does. Therefore, whilst covariance in power does indicate inter-brain coupling to a certain degree, it is not conclusive as EEG data have multiple dimensions not limited to power, and thus, the power covariance most often only indicates a weak form of association (Burgess, 2013). PLV and PDC focus on 'phases' of EEG data, which represent the relationship between the time and frequency domains of the data (Thatcher, 2012). PLV, as an example of phase synchrony, measures the degree of phase alignment specifically in the frequency domain (Lachaux et al., 1999), whilst PDC is designed to assess the direction of the alignment, where one source is assumed to have influence on the other (Baccalá & Sameshima, 2001). Yet, neither measure can suggest how the alignment occurs, as 'synchronisation' in these measures simply means that there is a consistent similarity or difference between two time-series data, regardless of the process of it emerging. The same applies to the measures commonly used in the hyperscanning fNIRS literature, which include wavelet transformed coherence (WTC) (e.g. Cui et al., 2012; Dommer, Jäger, Scholkmann, Wolf, & Holper, 2012; Holper, Scholkmann, & Wolf, 2012). WTC assesses the cross-correlation (or lagged correlation) between two time-series datasets as a function of frequency and time (Torrence & Compo, 1998), and therefore can evaluate 'turn-taking' patterns of relationship (Grinsted, Moore, & Jevrejeva, 2004). Nevertheless, similar to other coupling measures used in the fMRI and EEG literature, the analysis does not concern how the observed relationship between two datasets occurs, making it difficult to interpret the observed coupling activity between two individuals in relation to social or joint behaviour without other analyses.

Whilst coupling measures used so far simply represent the degree of synchrony between two time-series data without concerning the process of how the coupling occurred, a higher degree of inter-brain coupling in hyperscanning research is often interpreted as a higher level of temporal alignment of two data *due to* the reciprocal nature of interactions. Such interpretation is beyond what these indexes can actually capture and could be misleading, as synchronisation of two signals can be driven by various factors which are not necessarily related to social interaction. For instance, it has been shown that individuals who are presented with the same stimuli (e.g., a movie) typically exhibit a similar neural activation pattern, reflecting common information processing across individuals (Hasson et al., 2008; Hasson, Nir, Levy, Fuhrmann, & Malach, 2004). It is thus possible for participants who engage in the same data acquisition session to show similar neural activity, not because they are interacting with one another, but because they are in the same experimental setup. More importantly, it is also possible for synchronised activities to coincidentally emerge between individuals. For example, we can measure the EEG of two adults resting in separate rooms, both of whom would likely exhibit dominant alpha oscillatory activity. Even without any social interaction between them, we could expect to observe a consistent phase relationship between the two signals. These alternative scenarios where inter-brain coupling can emerge regardless of social interactional contexts point to the importance of evaluating to what extent the observed inter-brain coupling can be attributed to a social factor of interest, rather than a coincidence or a shared information processing. Moreover, a recent study demonstrated that these common measures could overestimate the strength of inter-brain coupling, failing to take into account that signal synchronisation can happen by chance (Burgess et al., 2013). This suggests that researchers' choice of measures for inter-brain coupling can bias the results, as well as the conclusion drawn from it. These suggest that we need to carefully examine how the reported inter-brain coupling is measured in each study to appropriately interpret the results.

It has also been suggested that certain study designs and procedures might be more likely to find a larger degree of inter-brain coupling than others, further pointing to the importance of examining the study design and procedure used in detail when discussing inter-brain coupling. For instance, based on the findings from research on mice, Kingsbury and colleagues (2019) proposed the mutual prediction theory (Kingsbury et al., 2019). Social interactions require participating individuals to act in accordance with predictions of their partner's behaviour. Each individual has brain systems that control their own behaviour, and also brain systems to monitor and predict their partner's behaviour. If these systems are localised in close proximity, neural activity of both systems are summed up when creating a measure to capture overall activities. If this is the case, a coherent pattern across multiple brains should arise. If this mutual prediction theory is accurate, some experimental setups might be more likely to find a higher degree of interpersonal connectivity than others. For example, generally, cooperative situations would require activations of brain systems monitoring their own behaviour as well as systems predicting other people's behaviour to a higher degree as compared to other situations such as competition (Czeszumski et al., 2021).

Without a systematic review, it is challenging to assess the effect of study designs on reported findings, as there is currently a large variation in published methods and analytic procedures across studies. Hyperscanning techniques have been used with different paradigms and tasks designed to assess different cognitive functions, such as decision making via economic games (e.g., Astolfi et al., 2011; King-Casas et al., 2005), successful communication and information transfer (e.g., Jiang et al., 2012), action coordination (e.g., Cui et al., 2012; Tognoli et al., 2007) or shared attention via mutual gaze tasks (e.g., Saito et al., 2010). Task setups for the same topic can also largely differ across studies, making it challenging to compare the same task. Some studies compared human dyads with other human dyads, some with human-computer dyads, or without a partner (e.g. Astolfi et al., 2020; Kourtis, Sebanz, & Knoblich, 2013; Osaka et al., 2015). It is possible that each parameter of these affects the reported results differently.

Not only the experimental settings but also participant characteristics might modulate the degree of inter-brain coupling. For instance, gender composition of dyads has been shown to modulate human social behaviour such as cooperation (Balliet, Li, Macfarlan, & Van Vugt, 2011), and indeed, recent studies have reported male-male, female-female or female-male dyads exhibit different levels of inter-brain coupling (Baker et al., 2016; Cheng et al., 2015; Li et al., 2020). There is no consensus so far as to which gender dyads display a stronger coupling, and if so, why there might be a difference. However, it is possible that gender composition of dyads could create a potential confound, and therefore, it is important to examine the current evidence systematically. Relationship or social closeness between interacting partners may also be a modulator of inter-brain coupling. Socially close individuals, such as romantic couples or parent-child dyads, experience many and rich opportunities to cooperate (Ackerman & Kenrick, 2009) or co-regulate their emotions (Feldman, 2007; 2015; Randal et al., 2013). Attachment formed through such a relationship also affects biological functions, such as the production of oxytocin (Feldman, 2007; Schneiderman, Zagoory-Sharon, Leckman, & Feldman, 2012). It is thus possible that a socially close relationship can modulate neural activity patterns via social behaviour and biological functions. Accordingly, some studies have investigated differences in inter-brain coupling across dyads in different relationships, and reported that socially close individuals show more similar patterns of neural activities during a social interaction (e.g., Pan, Cheng, Zhang, Li, & Hu, 2017a; Reindl et al., 2018). Despite such evidence, these participant characteristics are not always considered and controlled for. This can potentially contribute to a confound, making it challenging to interpret the degree to which observed interpersonal neural connectivity is associated with social interactions in the absence of other factors. Therefore, the effect of these participant characteristics on the reported inter-brain coupling is also worth examining in detail.

With the field reaching adequate numbers of empirical papers using hyperscanning techniques, it is timely to systematically review the existing literature, and explore whether, and if so how, variations in methodological choices might have influenced the reported results regarding inter-brain coupling that emerges during social interactions. Based on prior reviews and commentaries on hyperscanning studies discussed above (e.g., Czeszumski et al., 2020), the current meta-analytic review aims to assess methodological factors that fall into three categories: participant variables, task variables and analysis variables. If studies using certain methods are more likely to find a larger effect of inter-brain coupling than studies using other methods, that would help us appropriately interpret the existing evidence and better understand how interacting individuals influence each other at a neural level during a social interactional situation. There have been scoping review papers that have provided an overview of the characteristics of studies, pointing out gaps and issues in the existing literature (e.g., Babiloni & Astolfi, 2014; Czeszumski et al., 2020; Redcay & Schilbach, 2019). Yet, a systematic review involving extensive and systematic literature search is important to reduce the likelihood of bias involved in the review process, ensuring that we identify a comprehensive body of knowledge. In addition, a meta-analysis can indicate how robustly inter-brain coupling can occur during social exchanges.

Methods

This systematic meta-analysis was conducted based on the *a priori* protocol and the analysis plan formulated according to the NIRO framework (Topor et al., 2020). These were pre-registered (https://osf.io/9f2rg/) as per the practice guidelines offered by the Cochrane Collaboration (Higgins & Green, 2011).

Literature search and screening

Primary review questions

The primary review question of this meta-analytic review was regarding what methodological factors in hyperscanning studies might contribute to variations in reported findings regarding inter-brain coupling emerging during social interactions. To address this, the meta-analysis involved two steps. First, we evaluated the heterogeneity of the effect sizes reported in published studies discussing inter-brain coupling in a social setting across studies with different designs and methodologies. The second step of the meta-analysis explored how each methodological factor might influence the variation in the effect sizes reported. For this, the reported effect sizes were the dependent variable of the meta-analysis, whereas the methodological factors extracted from each study were independent variables.

Search strategy

A systematic online search as well as an additional bibliographic search (see “Alternative search” below) were undertaken in August 2020. No updating search has been undertaken.

Search terms. A series of searches using different sets of search terms was performed to locate as many relevant papers as possible, while minimising the number of irrelevant papers. The concepts for each search are illustrated in Figures 2-1, 2-2, and 2-3. The complete search strings are provided in Supplementary Material (Table S1). The search terms were finalised upon consulting a qualified librarian at Lancaster University. The process of finalising the search terms was guided by a tutorial provided by Staaks (2020). Search was performed in the title and abstract field, as recommended by a qualified librarian at Lancaster University.

Figure 2-1. Overview of search concept S1, S2, S3 and S4.

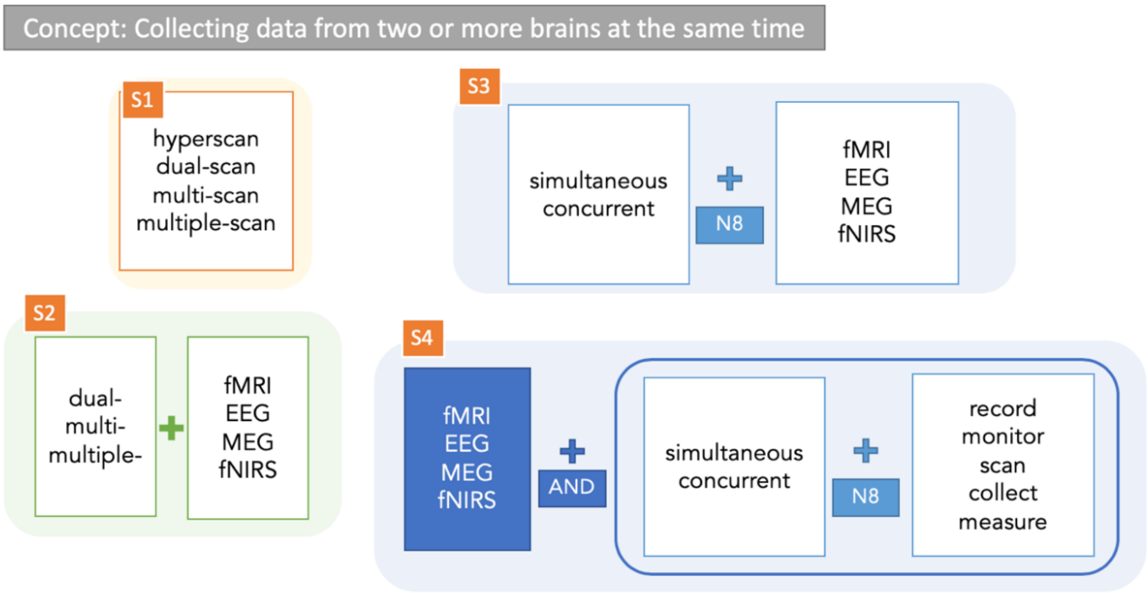


Figure 2-2. Overview of search concept S5.

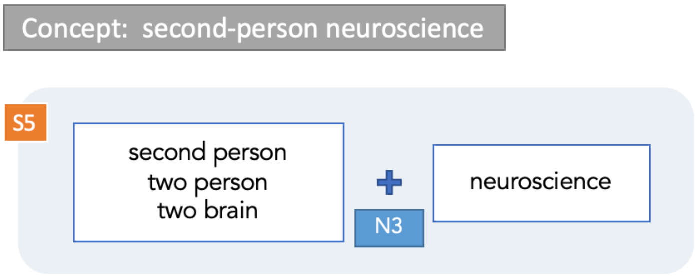
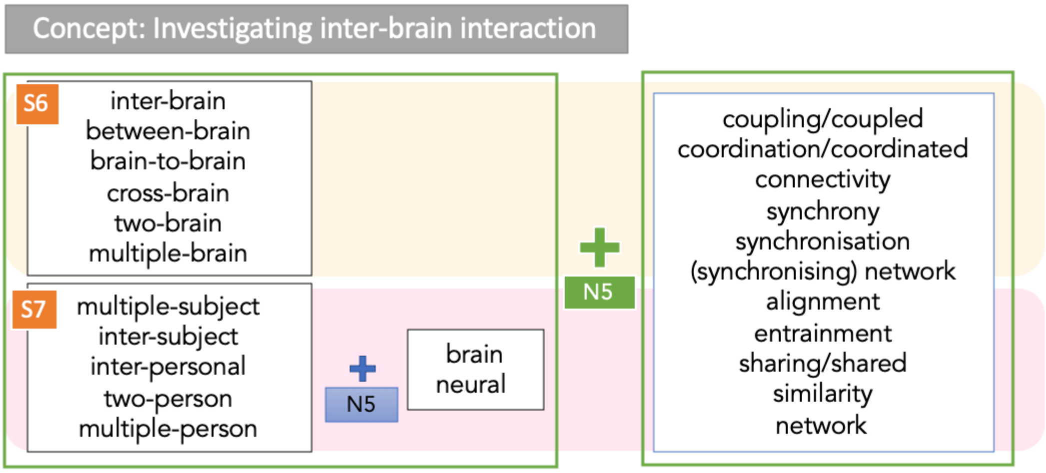


Figure 2-3. Overview of search concept S6 and S7.



Search limit. Search was limited to peer-reviewed articles published or accepted for publication in English between January 2002 and August 2020. The time frame was decided as the term hyperscanning was first introduced by Montague et al. in 2002. This review excluded any unpublished work and grey literature (i.e., literature produced outside of traditional commercial publishing and distribution channels) to control for the quality of studies. The hyperscanning technique is relatively new and still developing and there has been no standardised procedure established and validated. We expected that the limitation of papers to only papers published on peer-reviewed journals would effectively filter out studies of which experimental design, procedure or analysis was not optimal. Yet, excluding unpublished studies can increase a publication bias (Borenstein, Hedges, Higgins, & Rothstein, 2009). This was taken into account upon discussing the results of the review.

Database selection. Search was performed on APA PsycINFO (accessed via EBSCO), MEDLINE Complete (via EBSCO) and Scopus (via Elsevier). The selection was made upon consulting a qualified librarian at Lancaster University.

Alternative search. As our scoping search identified a large heterogeneity in the terms used in hyperscanning research, bibliographic search was also conducted in addition to electronic database search. This was to ensure relevant papers to the topic should be all included in the final set of papers for the subsequent data extraction and meta-analysis. During the process, the reference lists of ‘key papers’ were examined. These papers were selected upon consultation with two other colleagues who had at least published two papers using hyperscanning techniques at the time of this alternative search (2020). The list of these papers is provided in Supplementary Material (List S1). Furthermore, papers that cited one of these key papers were also screened. Any papers that were identified through this process but not located by the electronic search were included at this stage. These types of searches have been shown as valid and useful methods to locate studies as part of a systematic literature search (Brettle & Long, 2001; Hinde & Spackman, 2014; McNally & Alborz, 2004; Papaioannou, Sutton, Carroll, Booth, & Wong, 2010).

Criteria for inclusion and exclusion

The purpose of this review was to systematically describe the methodological variations in the existing literature, and quantitatively evaluate the effect of such variations on the variability of reported findings. Therefore, it was considered most optimal to include studies which might vary in methodology as long as the research question concerns inter-brain coupling based on the data were collected from multiple different individuals. Moreover, as mentioned above, the existing review papers have identified a large variability in methods in hyperscanning studies. Hence, it was likely that focusing on one type of studies could make the number of studies to be included too small to capture the overall picture of the field. Setting relatively flexible inclusion criteria was expected to allow us to retain the wholeness of the existing studies as a body of literature, while preserving the idiosyncratic nature of each study. The criteria used for screening are described below.

The main effect of interest. Research assessing the association between two people's brain activities during a social or joint task were included. Studies were included as long as there was at least one interactional factor between participants, even if it might be via screen, and if data were acquired in different timings and places (e.g. Leong et al., 2017).

Participant groups of interest. Only research involving human participants were included. Research involving non-human samples were excluded.

Study designs. Only experimental studies were included. Theoretical, methodological or review papers were excluded.

Data reporting. Studies were included when it reported sufficient information to allow the calculations of effect sizes and their variances, or the authors provided additional information upon request. When the effect size was not reported, original data or information such as *N*s, *t*-values or *F*-values and *p*-values were requested to the corresponding author of the paper to enable the calculation of the effect sizes. Studies reporting no statistics were excluded.

Screening

The screening was conducted primarily by one reviewer according to the PRISMA guideline. For title and abstract screening, search results were downloaded from each database as csv. files, and uploaded on Rayyan, an online screening programme. Using the criteria described above, titles and abstracts were screened, and potentially relevant articles (*n*=276) were identified. For full-text screening, the remaining articles were imported into Mendeley (a bibliography management software) and each paper was marked as include or exclude. The number of articles marked as included was *n*=135. The screening results were manually transferred to a Microsoft Excel spreadsheet, which was then used for the subsequent data extraction.

Extracting and coding research for the review and meta-analysis

Extracting the effect sizes

Effect sizes of significant main effects for inter-brain coupling were extracted from the papers identified as eligible via full-text screening. When multiple-group comparisons were made, effect sizes from post-hoc comparisons of two groups were extracted instead of an effect size from an omnibus test (Borenstein et al., 2009).

Effect sizes reported as non-significant based on a frequentist inference (i.e., *p*-value greater than .05) were not included in the meta-analysis. This was because the focus of the current meta-analysis was to examine potential influence of methodological variations on reported effect sizes in the existing literature, not to provide an accurate estimate of overall effect size of inter-brain coupling. The key question which the current meta-analysis aimed to address was whether ‘significant’ brain-to-brain activity that has been reported to occur during social interactions could be a methodological artefact or a coincidental resemblance due to other reasons but social factors (Burgess et al., 2013, Hamilton, 2021; Hasson et al., 2008). Hence, it was considered as appropriate to only focus on significant results. Whilst this limitation is likely to lead to the overestimation of the overall effect sizes given that the effect sizes of non-significant results are likely to be smaller than the effect sizes of significant results (Feng, Thompson, & Paulus, 2022), this should not affect our analysis on potential influence of methodological variations in reported effect sizes in published reports. Whilst conventional meta-analyses have included both significant and non-significant results with the view to provide an accurate overall effect size from the comprehensive sources of literature, the approach of targeting sources to be reviewed to published and/or significant effects have been done in other meta-analyses (e.g., Siemens et al., 2021), and this has been an effective method to investigate methodological biases that are present even among the most rigorously controlled sources of scientific evidence (i.e., published effects that are statistically examined and found to be significant) (Howick et al., 2022; Siemens et al., 2021). As the aim of the current meta-analysis was not to accurately estimate the overall effect size, the direction of the effect size was not considered, and all the effect sizes were extracted as absolute values.

It was a common practice across the eligible studies that researchers conducted additional analyses when the main analysis found a significant result. In such cases, we only included the results of the main analyses that were directly related to the main research question of the paper. Accordingly, additional analyses following the main analyses were excluded. When the authors conducted multiple comparisons on each pair of channels, voxels or clusters, effect sizes calculated from the pairs which showed the significant effect were included and effect sizes from the pairs which showed no significant effect were excluded. It was also a common procedure, especially in fNIRS studies, that researchers compare empirical data with surrogate data to identify channel pairs that are more likely to show the effect relevant to the task used, or to examine the effect observed is more significant than a chance level (i.e., validation analysis). Effect sizes from these results were excluded when these comparisons between actual and surrogate data were followed by a main analysis such as a comparison between different experimental conditions. If the comparison between actual and surrogate data was treated as the main analysis of the study, the effect sizes from these comparisons were included. Relatedly, particularly in fMRI papers, it is common for authors to report cluster-level inferences (i.e., a group of neighbouring voxels) and voxel-level inferences. For a cluster-level analysis, it was conventional to only report *p*-values, whereas both *t*-values and *p*-values were often reported for a voxel-level analysis. In this case, relevant statistical information was requested via email to the corresponding author. When the information was not available, the results of cluster-level analysis were excluded from the current meta-analysis.

Calculating the effect sizes and its standard error

When the effect sizes and its standard error (*SE)* were not reported in the original paper, the estimates were calculated based on the reported statistics, including *F*-, *p*-, and *t*-values, using the calculator distributed online (Lakens, 2013) as well as R dmetar package (Harrer, Cuijpers, Furukawa, & Ebert, 2019). When additional information was required to calculate the effect size estimate, the corresponding author of the article was contacted via email. If no reply was received within four weeks after the initial contact, a reminder email was sent. If there was still no reply for two weeks after the reminder, the study was excluded due to the missing information, because it was impossible to calculate the effect size estimate.

To enable the conversion of effect sizes to Cohen's *d* at a later stage of this meta-analysis, when the partial eta-squared was reported as an effect size, the eta-squared was calculated using the spreadsheet by Laken (2013), based on the data reported on the manuscript. When the Spearman's correlation coefficient was reported, although it is technically and qualitatively different from Pearson's correlation coefficient, it was treated equally as a Pearson's correlation coefficient when converting it to Cohen's *d*. When a non-parametric test was used, such as Wilcoxon signed-rank test, a correlation coefficient was estimated using the equation of $r = Z/N^2$ (If paired, *N* is typically the number of pairs) (Pallant, 2010).

Coding of literature

The PDFs of published articles selected via full-text screening were coded according to the coding scheme, which are detailed below in this section. When the required information was missing in a published manuscript, an email was sent to a corresponding author of the paper. If no additional information was obtained, the data for the analysis was coded as "not reported".

(A) Participant variables

Sample size. The final sample size that contributed the data to the final analysis was extracted. When it was unclear whether the reported sample size was the final sample or recruited sample, the reported sample size was extracted.

Participant age. The participant age coding included (1) “adults”, referring to dyads consisting of two adults and (2) “adult-child”, referring to dyads consisting of an adult (aged over 19) and child (aged under 18). The grouping was done according to the average age of participants reported. When the mean age of the recruited participants was reported, but not the subgroup (or final sample) whose data were used for the analysis, the reported mean age of the originally recruited group was extracted. When there were two roles in participants (e.g., students and teachers) and the average age was reported per role, the mean age of the group which composed of the larger portion of the whole sample was extracted (e.g., For Zhang et al., (2020), where the student group consisted of *n*=60 whereas the teacher group was composed of *n*=4, the mean age of the student group was extracted)

Gender composition of participant dyads. The participant gender composition coding included (1) “same”, referring to dyads consisting of individuals of the same gender (i.e., female-female and/or male-male), (2) “opposite”, referring to dyads consisting of individuals of the opposite gender (i.e., female-male), (3) “mixed”, referring to the case where the gender was not controlled in making a dyad and hence there were both opposite-gender and same-gender dyads, (4) “manipulated”, referring to the case where the aim of the analysis was to assess the difference between different types of gender composition, and (5) “not reported”.

Relationship of participant dyads. The coding for the dyad relationship involved (1) “strangers”, referring to dyads who had not known each other well prior to the data collection, and (2) “close”, referring to dyads who had known each other and had formed certain relationship, such as family, romantic couples, colleagues in the same institution, and friends. Acquaintances were categorised as “strangers”, as these were often used to describe dyads who had only seen each other in a classroom before, but were not friends. When the aim of the analysis was to assess the difference between different levels of social closeness, it was coded as (3) “manipulated”.

(B) Task variables

Interaction medium. Interaction medium refers to a form of interaction taking place. The coding for this included (1) “computer” (i.e., interaction via a computer-based task), (2) “direct” (i.e., face-to-face interaction), (3) “limited” (e.g., interaction with selected modality, such as interactions where participants could see each other but were not allowed to communicate verbally), (4) “mixed” (i.e., data of interest collected from multiple settings involving different media). When the aim of the analysis was to evaluate the difference between different types of interaction medium, it was coded as (5) “manipulated”.

Interaction modality. Interaction modality refers to a main type of sensory information conveyed during an interaction, or a type of stimuli used in a social experimental task. The coding included (1) “auditory stimuli only” (e.g., verbal interaction without seeing each other), (2) “visual stimuli only” (e.g., participants could see each other but were not allowed to verbally interact), (3) “visual and auditory stimuli” (e.g., naturalistic interaction), (4) “computer-based” (i.e., no in-person interaction). When the aim of the analysis was to evaluate the difference between different types of interaction modality, it was coded as (5) “manipulated”.

Cognitive function underlying the task. Based on the classification by Czeszumski et al. (2020), eligible analyses were categorised according to a cognitive function that was aimed to examine utilising a hyperscanning technique. The coding involved (1) “speech and communication”, (2) “competition vs cooperation”, (3) “action coordination”, (4) “game and decision-making”, (5) “action representation and joint gaze”, and (6) “membership”.

(C) Analysis variables

Measures of inter-brain coupling. Based on the categorisation by Czeszumski et al. (2020), eligible analyses were grouped according to the type of measures used to assess inter-brain coupling. The coding involved (1) “coupling/connectivity analysis” (e.g., Phase-Locking Value, Phase-Lag Index, phase coherence), (2) “correlation or dependence analysis” (e.g., partial or semi-partial correlation coefficient, total independence analysis), (3) “graph theory index” (e.g., modularity, density, small world-ness), and (4) “information flow” (e.g., Granger causality, Partial Directed Coherence).

Statistical comparisons. As the current review and meta-analysis involved effect sizes from a comparison between two experimental conditions as well as ones from a comparison between experimental and control conditions, we examined the influence of statistical assessment of choice in hyperscanning research on the reported effect. The coding included (1) “AB” (i.e., comparing between different experimental conditions), (2) “AC” (i.e., comparing between an experimental condition and a control condition), (3) "correlational" (i.e., only one experimental condition and correlation between factors measured in the same condition was assessed), (4) "chance level" (i.e., comparison with a chance level) and (5) "surrogate" (i.e., comparison between actual data and surrogate data).

Test of inter-rater reliability

All studies were primarily coded by one coder. To avoid potential bias by a single coder, a subset of studies (30% for title and abstract, 20% for full-text screening) was given to the total of three other coders (two for title and abstract screening, one for full-text screening). Each reviewer independently conducted screening according to the same coding scheme. The initial agreement rate between coders was 94% at the title and abstract stage, and 90% at the full-screening stage. When the coders' decisions were not in agreement, the discussion was held between the coders, during which a joint decision was made.

Meta-analytic methods

Pooling effect sizes

Cohen's *d*, representing a difference between the means of conditions, was chosen as a common metric in the meta-analysis. All the effect size extracted were converted to Cohen's *d* using an online effect size converter (https://www.escal.site/). To convert Hedges' *g* to Cohen's *d*, the asymptotical equation given by Hedges and Olkin (1985) was used. Eta-squared was converted to a correlation coefficient *r* (when the degree of freedom (*df*) = 1) or R^2 (when the *df* is larger than 1), which was further converted to Cohen's *d*. When the effect was reported to be significant without any further statistic, we assumed $p = .05$. When the effect was reported with "$p <.001$", we assumed $p =.001$. Similarly, when "$p <.005$" was reported, we assumed $p =.005$, and when "$p <.05$" was reported, we assumed $p =0.05$.

Given that the current meta-analysis examined multiple analyses (i.e., multiple effect sizes) extracted from one study, a multi-level meta-analysis was performed. The use of multilevel model was selected because effect sizes extracted from one study were assumed to correlate with one another. Such correlations were accounted in a multilevel meta-analysis model (Higgins & Green, 2011), which enabled us to assess whether accounting for from which paper an effect size was extracted better would explain the variability in the reported effect sizes (Cheung, 2019). The multilevel meta-analysis considered "Paper" as the level three, and "Analyses" as the level two. The analysis was conducted using the rma.mv function in the "metafor" package (Viechtbauer, 2010) in the R software environment (R Core Team, 2014). The multilevel meta-analysis provided an overall effect size estimate (Cohen's d) for inter-brain coupling across all the eligible analyses, a 95% confidence interval and an overall p value for d, as well as parameters for heterogeneity at the within- (σ^2_1) and between-studies (σ^2_2) levels. To assess the heterogeneity of all the studies included, I^2 was computed (Higgins & Thompson, 2002).

Subgroup analyses via meta-regression

The effect of methodological variables on the reported effect sizes was assessed using a meta-regression. The analysis was performed using an R function rma.mv in a metafor package (Viechtbauer, 2010), where each grouping variable was added to the model as a moderator. One of the levels in each variable was treated as a reference, and the results from the omnibus test as well as post-hoc comparisons were examined to test whether there was any difference between subgroups in terms of the overall effect size estimate. The overall effect size estimate of each subgroup was also calculated by fitting a meta-regression model without an intercept.

Assessment of small sample bias

The effect of the data censoring (i.e., bias) on the outcome of the meta-analysis was assessed using a funnel plot. A funnel plot illustrates an effect size of each analysis against the standard error (Borenstein et al., 2009). This plot having a symmetrical funnel shape indicates no data censoring towards analyses with a large sample. If the bottom left-hand corner of the plot is more scarce than the other corners of the plot, that suggests that the included dataset might be biased towards analyses with a larger sample, due to the inclusion of fewer analyses with a smaller sample. This is because studies with a small sample typically have a smaller statistical power, and hence are less likely to be published (Mullen, 2013).

Results

Systematic search

The systematic search identified 7268 provisionally eligible sources, which were then screened according to the title and abstract of each article. The title and abstract screening identified 276 relevant papers. With 47 papers from the alternative searches (e.g., bibliographic search) added, the total of 323 papers were subjected to full-text screening. The number of papers which met the inclusion criteria was *n*=133, whilst 190 papers were excluded due to the study design (*n*=134), the measure (*n*=5), the publication style (*n*=17), the content (review *n*= 3, theoretical *n*=9, methods *n*=22). Additional 39 papers were excluded from the meta-analysis because they did not report sufficient statistics to calculate the effect size (*n* =15), found no main effect (*n*= 23) or used a total number of channel/voxel/cluster pairs in each condition to investigate an inter-brain coupling (*n*=1). Furthermore, due to the small number of eligible papers using MEG (*n*=2), these MEG papers were also excluded. The final set of papers eligible for the meta-analysis was composed of 93 papers, including 23 EEG papers, 58 fNIRS papers, and 12 fMRI papers, which reported 132 EEG analyses, 229 fNIRS analyses, and 80 fMRI analyses. The results of the systematic search following the PRISMA-P guideline are also summarised in Figure S1 and Table S2 in Supplementary Material. The lists of papers included in the final meta-analysis are provided in Supplementary Material (Lists S2, S3, S4).

Meta-analysis results

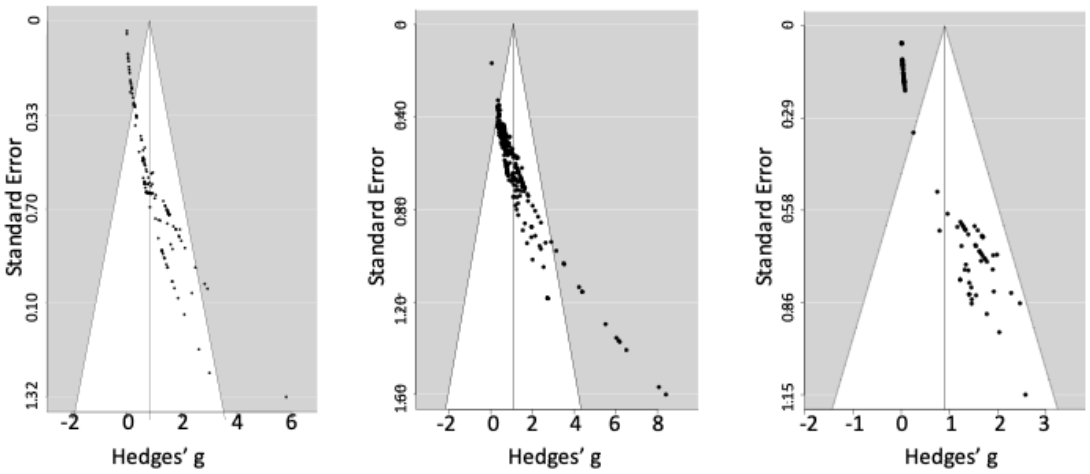
Overall effect for inter-brain coupling and heterogeneity analysis

A multi-level meta-analysis of inter-brain coupling across analyses involving EEG measures ($k = 132$) found a significant large effect (Cohen's $d = 0.86$, 95%*CI* [0.60, 1.11], $p < .001$, $p < .001$, $\sigma^2_1 = .29$, $\sigma^2_2 <.001$). Notably, a substantial heterogeneity (total I^2= 79.86%) was observed (Higgins, Thompson, Deeks, & Altman, 2003). This observed heterogeneity was largely attributed to a between-study variance (Analysis [Level2] I^2 <0.001, Study [level 3] I^2 =79.86%). A multi-level meta-analysis across analyses involving fNIRS measures (k = 229) found a significant large effect (Cohen's $d = 1.08$, 95%*CI* [0.91, 1.26], $p < .001$, $\sigma^2_1 = .30$, $\sigma^2_2 <.001$). There was a moderate heterogeneity (total I^2= 50.00%), which was largely attributed to a between-study variance (Analysis [Level2] I^2<0.001, Study [level 3] I^2 = 50.00%). A multi-level level meta-analysis across analyses involving fMRI measures (k = 80) found a significant large effect (Cohen's d = 0.94, 95%CI [0.46, 1.42], p = .002, σ^2_1 = .47, σ^2_2 <.001). A moderate heterogeneity (total I^2= 64.93%) was observed, which was mainly attributed to a between-study variance (Analysis [Level2] I^2<0.001, Study [level 3] I^2 = 64.93%).

Small sample bias analysis

The funnel plot on the eligible studies (Figure 2-4) indicated the current-meta analysis mainly included studies with a large sample size reporting a smaller effect as well as studies with a smaller sample reporting a larger effect. Since the current meta-analysis only included peer-reviewed and published studies, a certain level of publication bias was expected.

Figure 2-4. Funnel plots on the eligible EEG studies (left), fNIRS studies (middle) and fMRI studies (right).



Note. Y-axis shows the standard error, with larger standard errors plotted on the upper side. X-axis shows the estimated Hedge's g of each analysis included in the meta-analysis.

Analysis of the Impact of methodological variation

To calculate an estimate of overall effect size for each subgroup, meta-regression analyses without an intercept were performed, of which results are summarised in Table S3 in Supplementary Material. In the following, we report the results of the meta-regression analyses taking one of the subgroup as a reference to examine the difference across subgroups. The overall effect size estimate for each subgroup for each analysis is provided in Table S3 in Supplementary Material. Forest plots for each technique are also provided in Supplementary Material (Figures S5, S6 and S7).

Sample size

A meta-regression including sample size as a moderator on EEG analyses suggested that analyses involving a smaller sample size found a larger effect than analyses involving a larger sample size (*F*(1,130) = 4.24, *p* = .04, *QE*(130) = 266.18, *p* <.0001). Likewise, a meta-regression on fNIRS analyses indicated that analyses based on a smaller sample found a larger effect than ones based on a larger sample (*F*(1,227)= 5.69, *p* = .018, *QE*(227) = 321.89, *p* =.0001). On the other hand, a meta-regression on fMRI studies suggested no significant association between the effect sizes and the sample sizes (*F*(1,78) = .20, *p* = .65, *QE*(78) = 111.10, *p* =.0082).

Participant age group

Eligible analyses were grouped into two categories according to the mean age of the participants reported: (1) “adults” (aged over 19 years) and (2) “adult-child” (aged 18 or younger). The effect of participant age groups was evaluated by a three-level meta-regression adding participant age as a categorical moderator.

For the EEG studies, the meta-regression taking the “adult-child” as a reference suggested that participant age group did not significantly affect the reported effect sizes on inter-brain coupling (omnibus test: *F*(1,130)= 0.0014, *p*=.97, *QE*(130)=305.10, *p*<.0001). With the fNIRS studies, the meta-regression taking the “adult-child” as a reference indicated that participant age group did not significantly explain the variance in the reported effect sizes on inter-brain coupling (omnibus test: *F*(2,227) = 1.21, *p*=.27, *QE*(227)= 321.93, *p* <.0001). Regarding the fMRI studies, the meta-regression taking the “adult-child” as a reference suggested that participant age group did not significantly affect the reported effect sizes on inter-brain coupling (omnibus test: *F*(2,76) =1.07, *p* =.37, *QE*(76)= 200.57, *p* <.0001).

Participant gender composition

The effect of gender composition of the dyad on the reported effect sizes was assessed by a meta-regression adding Gender Composition as a categorical moderating factor. There were five subgroups of Gender Composition: (1) “opposite” (male (M) – female (F)), (2) “same” (MM, FF or a mix of MM and FF), (3) “mixed” (i.e., gender composition was not controlled), (4) “manipulated” (i.e., different gender dyads were compared), and (5) “not reported”.

For the EEG studies, the meta-regression taking the “opposite” as a reference indicated that gender composition did not significantly influence the reported effect sizes on inter-brain coupling (omnibus test: $F(3,128) = 1.064$, $p = .37$, $QE(128) = 253.37$, $p < .0001$). With the fNIRS studies, the meta-regression taking the “opposite” as a reference suggested that gender composition had a significant influence on the reported effect sizes on inter-brain coupling (omnibus test: $F(2,224) = 3.30$, $p = .012$, $QE(224) = 284.07$, $p = .004$). Post-hoc comparisons were conducted by examining the difference between a reference subgroup and each of the other subgroups (see Table 2-1), as well as comparing regression coefficients within the meta-regression models (see Table 2-2). These post-hoc analyses suggested that the subgroup of analyses explored inter-brain coupling in opposite gender dyads found an overall larger effect size than other subgroups. Yet, the sample size of the subgroup of analyses that examined inter-brain coupling between opposite gender dyads was $k = 15$, which was from three studies, smaller than the number of analyses included in other subgroups. This unequal sample size needs to be considered when interpreting this result. Regarding the fMRI studies, there were no analyses involving dyads of opposite genders (i.e., “opposite”) or ones comparing dyads of different gender compositions (i.e., “manipulated”). Hence, the meta-regression assessed the variance component in three subgroups. The meta-regression taking the “same” as a reference indicated that gender composition did not significantly affect the reported effect sizes on inter-brain coupling (omnibus test: $F(2,77) = 1.19$, $p = .31$, $QE(77) = 191.71$, $p < .0001$).

Table 2-1.Results of the post-hoc comparisons following the meta-regression taking the "opposite" as a reference.

	Estimate	*SE*	*t*-value	*p*-value
Same	-0.98	.36	-2.70	.0074*
Manipulated	-1.14	.51	-2.24	.0261*
Mixed	-1.22	.35	-3.49	.0006*
Not reported	-0.98	.36	-2.70	.0074*

Table 2-2. Results of the post-hoc comparisons following the meta-regression taking the "opposite" as a reference, by comparing across regression coefficients within the meta-regression model.

	F-value	*p*-value
Manipulated vs Same	3.31	.0117*
Manipulated vs Mixed	6.09	.0027*
Manipulated vs Not reported	4.36	.0052
Mixed vs Same	4.36	.0052
Mixed vs Not reported	6.49	.0018*
Not reported vs Same	3.65	.0275*

Participant Relationship

The effect of participant relationship within a dyad on the reported effect sizes was assessed by a meta-regression adding Relationship as a categorical moderating factor. There were five subgroups of Relationship: (1) "strangers" (e.g., strangers or acquaintances but not in a specific relationship), (2) "close" (e.g., family members, friends, or colleagues), (3) "mixed" (i.e., participant relationship was not controlled), (4) "manipulated" (i.e., stranger dyads and non-stranger dyads were compared), and (5) "not reported".

With the EEG studies, there were no analyses comparing stranger dyads and non-stranger dyads (i.e., "manipulated"), nor analyses that did not control participant relationship (i.e., "mixed"). Hence, the meta-regression was performed across three subgroups. The meta-regression taking the "strangers" as a reference suggested that participant relationship within a dyad was not significantly related to the reported effect sizes on inter-brain coupling (omnibus test: $F(3,128) = .3522$, $p = .79$, $QE(128) = 292.33$, $p < .001$). Regarding the fNIRS studies, the meta-regression taking the "strangers" as a reference indicated that participant relationship within a dyad did not significantly influence the reported effect sizes on inter-brain coupling (omnibus test: $F(4, 224) = 1.62$, $p = .171$, $QE(225) = 318.57$, $p < .001$). For the fMRI studies, there were no analyses comparing stranger dyads and non-stranger dyads (i.e., “manipulated”, nor analyses that did not control participant relationship (i.e., “mixed”). Hence, the meta-regression was performed across three subgroups. The meta-regression taking the “strangers” as a reference suggested that participant relationship within a dyad did not significantly affect the reported effect sizes on inter-brain coupling (omnibus test: $F(2,77) = .85$, $p = .43$, $QE(77) = 199.94$, $p < .0001$).

Interaction medium

The effect of interaction medium on the reported effect sizes was assessed by a meta-regression adding Medium as a categorical moderating factor. There were five subgroups of Medium: (1) “computer” (i.e., interaction via a computer-based task), (2) “direct” (i.e., face-to-face interaction), (3) “limited” (i.e., interaction with selected modality; e.g., interactions where participants could see each other but were not allowed to communicate verbally), (4) “mixed” (i.e., analyses of data from multiple settings involving different media) and (5) “manipulated” (i.e., analyses comparing interactions with different interactional media). Due to the nature of fMRI data acquisitions, all the eligible fMRI studies in this meta-analysis utilised a screen-based interaction. Hence, the meta-regression exploring the effect of Medium was not performed for the fMRI studies.

With the EEG studies, there were no analyses involving data from the settings including different interaction media (i.e., “mixed”). Thus, the meta-regression was performed across four subgroups. The meta-regression taking “direct” as a reference suggested that interaction medium did not significantly affect the reported effect sizes on inter-brain coupling (omnibus test: $F(3,128) = .52$, $p = .67$, $QE(128) = 241.62$, $p < .001$). Regarding the fNIRS studies, the meta-regression taking “direct” as a reference indicated that interaction medium significantly influenced the reported effect sizes on inter-brain coupling (omnibus test: $F(3,225) = 4.41$, $p = .0049$, $QE(225) = 277.34$, $p = .0010$). Post-hoc comparisons were conducted by examining the difference between a reference subgroup and each of the other subgroups (see Table 2-3), as well as comparing regression coefficients within the meta-regression model (see Table 2-4). This post-hoc analysis suggested that the subgroup of analyses explored inter-brain coupling emerging during an interaction of “limited” medium found an overall larger effect size than other subgroups. Nonetheless, the number of analyses included in this “limited” subgroup was *k*= 8, all of which were reported in the same study (*n*=1). Hence, this result needs to be interpreted with caution.

Table 2-3. Results of the post-hoc comparisons following the meta-regression taking the “direct” as a reference.

	Estimate	*SE*	*t*-value	*p*-value
Computer	0.14	0.20	0.74	.46
Limited	2.14	0.60	3.54	.0005*
Manipulated	-0.08	0.24	0.72	.72

Table 2-4. Results of the post-hoc analysis that compared across regression coefficients within the meta-regression model.

	F-value	*p*-value
Computer vs Limited	6.43	.0019*
Computer vs Manipulated	4.51	.0043*
Limited vs Manipulated	6.58	.0017*

Interaction modality

The effect of interaction modality on the reported effect sizes was assessed by a meta-regression adding Modality as a categorical moderating factor.

With regards to the EEG studies, there were five subgroups of Modality: (1) “visual stimuli only” (e.g., participants could see each other but were not allowed to verbally interact), (2) “auditory stimuli only” (e.g., verbal interaction without seeing each other) , (3) “visual and auditory stimuli” (e.g., naturalistic interaction), (4) “computer-based” (i.e., no in-person face-to-face interaction), and (5) “manipulated” (i.e., analyses comparing interactions with different modalities). The meta-regression taking the “visual only” as a reference indicated that interaction modality did not significantly affect the reported effect sizes on inter-brain coupling (omnibus test: $F(4,127) = .43$, $p=.79$, $QE(127)= 208.81$, $p < .0001$). With the fNIRS studies, there were five subgroups of Modality: (1) “visual stimuli only” (2) “auditory stimuli only”, (3) “visual and auditory stimuli”, (4) “computer-based”, and (5) “manipulated”. The meta-regression taking the “visual only” as a reference indicated that interaction modality did not significantly influence the reported effect sizes on inter-brain coupling (omnibus test: $F(4,224)= .30$, $p = .88$, $QE(224)= 316.62$, p

<.0001). For the fMRI studies, there were four subgroups of Modality: (1) “visual stimuli only”, (2) “auditory stimuli only”, (3) “motor task” (e.g., hand gripping), and (4) “visual stimuli and motor task” (e.g., model origami folding). The meta-regression taking the “visual only” as a reference indicated that interaction modality had no significant variance component on the reported effect sizes on inter-brain coupling (omnibus test: $F(3,76)= .93$, $p =.43$, $QE(76)= 193.12$, $p <.0001$).

Cognitive function underlying the task

The effect of cognitive function underlying the experimental task on the reported effect sizes was assessed by a meta-regression adding Cognition as a categorical moderating factor.

With regard to the EEG analyses, there were six subgroups: (1) “action coordination”, (2) “competition vs cooperation”, (3) “game and decision-making”, (4) “speech and communication”, (5) “membership”, and (6) “action representation and joint gaze”. The meta-regression taking the “action coordination” as a reference indicated that an experimental task choice did not significantly affect the reported effect sizes on inter-brain coupling (omnibus test: $F(5,126)= .28$, $p =.92$, $QE(126)= 222.78$, $p <.0001$). For the fNIRS analyses, six subgroups were identified: (1) “action coordination”, (2) “competition vs cooperation”, (3) “game and decision-making”, (4) “speech and communication”, (5) “action representation and joint gaze”, and (6) “manipulated” (i.e., analyses comparing different tasks). The meta-regression taking the “action coordination” as a reference indicated that an experimental task choice did not significantly affect the reported effect sizes on inter-brain coupling (omnibus test: $F(5,223)= .97$, $p =.44$, $QE(223)= 294.14$, $p =.001$). Regarding the fMRI analyses, there were five subgroups identified: (1) “action coordination”, (2) “competition vs cooperation”, (3) “game and decision-making”, (4) “speech and communication”, (5) “action representation and joint gaze”. The meta-regression examining the variance components in five subgroups taking the “action coordination” as a reference indicated that an experimental task choice did not significantly impact the reported effect sizes on inter-brain coupling (omnibus test: $F(4,75)= 1.83$, $p =.13$, $QE(75)= 95.65$, $p =.054$).

Measure of inter-brain coupling

The effect of measures used to assess inter-brain on the reported effect sizes was assessed by a meta-regression adding Measure as a categorical moderating factor.

Regarding the EEG analyses, the measures for inter-brain coupling were grouped into four subgroups: (1) “coupling or coherence measure” (e.g., Phase Lag Value, Phase Lag Index, phase coherence), (2) “correlation” (e.g., Pearson’s correlation, Spearman’s correlation), (3) “information flow” (e.g., Granger causality, Partial Directed Coherence) and (4) “graph index” (e.g., density, small-worldness). The meta-regression taking the “coupling” as a reference indicated that choices in the measures did not have significant influence on the reported effect sizes on inter-brain coupling (omnibus test: $F(3,128) = .19$, $p = .90$, $QE(128) = 223.55$, $p < .0001$). For the fNIRS analyses, measures for inter-brain coupling were grouped into four subgroups: (1) “coupling or coherence measure” (e.g., Wavelet Transform Coherence), (2) “correlation” (e.g., Pearson’s correlation, Spearman’s correlation), (3) “information flow” (e.g., Granger causality, Partial Directed Coherence) and (4) “psychophysiological interaction analysis” (PPI). The meta-regression taking the “coupling” as a reference indicated that measures chosen did not significantly influence the reported effect sizes on inter-brain coupling (omnibus test: $F(3,225) = .29$, $p = .83$, $QE(223) = 314.38$, $p = .001$). With regards to the fMRI analyses, measures for inter-brain coupling were grouped into two subgroups: (1) “correlation” and (2) “information flow”. The meta-regression taking the “corelation” as a reference indicated that choices in the measures did not explain the variance in the reported effect sizes on inter-brain coupling (omnibus test: $F(1,78) = .62$, $p = .46$, $QE(78) = 194.26$, $p < .0001$).

Statistical comparison

The effect of a statistical comparison performed was examined using a three-level meta-regression taking Statistical Comparisons as a moderating categorical variable. Statistical comparisons were coded as follows: (1) “AB” (i.e., comparing between different experimental conditions), (2) “AC” (i.e., comparing between an experimental condition and a control condition), (3) “correlational” (i.e., only one experimental condition and correlation between factors measured in the same condition was assessed), (4) “chance level” (i.e., comparison with a chance level) and (5) “surrogate” (i.e., comparison between actual data and surrogate data).

For the EEG analyses, there were no analyses in the "chance level" subgroup. Hence, the meta-regression was performed on four subgroups. The meta-regression taking the "AB" as a reference indicated that types of statistical comparison did not have a significant influence on the reported effect sizes on inter-brain coupling (omnibus test: $F(3,128) = .78$, $p = .51$, $QE(128) = 207.05$, $p < .0001$). With the fNIRS analyses, there were no analyses in the correlational subgroup. Hence, the meta-regression was performed on four subgroups. The meta-regression taking the "AB" as a reference indicated that types of statistical comparison did not significantly influence the reported effect sizes on inter-brain coupling (omnibus test: $F(3,225) = 1.49$, $p = .22$, $QE(225) = 325.92$, $p < .0001$). With regards to the fMRI analyses, there were no analyses in the "correlational" subgroup. Hence, the meta-regression was performed on four subgroups. The meta-regression taking the "AB" indicated that types of statistical comparison did not significantly influence the reported effect sizes on inter-brain coupling (omnibus test: $F(3,76) = 2.33$, $p = .08$, $QE(76) = 51.41$, $p = .99$).

Discussion

Investigating the relationship between the brains of interacting individuals utilising hyperscanning techniques has a large potential to advance our understanding of neural mechanisms underlying social reciprocal interactions. Yet, cautions have been raised in interpreting the results of these hyperscanning studies, as the observed 'synchronisation' between two brain activities could be, at least in part, due to methodological artefacts (Burgess, 2013; Hamilton, 2021). To better understand the impact of methodological choices on the reported inter-brain coupling (or 'synchronised' pattern seen across different brains) in a social situation or during a joint activity, we systematically evaluated the heterogeneity of the results published thus far, and examined how these results might differ depending on certain methodological factors. The systematic search, on which the meta-analysis was based, identified 23 EEG papers involving 132 eligible analyses for the current meta-analysis, 58 fNIRS papers including 229 analyses, and 12 fMRI involving 80 analyses.

With regards to our first aim of assessing the variability in the reported effect sizes across published studies, our meta-analysis indicated a moderate to substantial heterogeneity in the effect sizes reported across studies. Our analysis suggested that, regardless of the brain imagining techniques used, there was a moderate to substantial heterogeneity (50.00 - 79.86%). Importantly, the observed heterogeneity was consistently largely associated with a between-study variability rather than between-analysis variability. This indicates that effect sizes extracted from the same study were similar to one another within the same study, whereas the overall effect sizes of each study differ to a large extent from one another.

Given the large variation in the reported effect sizes across different methodological approaches, we further examined methodological factors that might potentially contribute to this observed heterogeneity in the reported effect sizes. A total of nine potential methodological variables were examined as a potential moderator of the reported effect sizes. Meta-regressions examining EEG and fMRI suggested that the observed heterogeneity was not accounted for by any of these specific methodological variables. Meta-regressions examining fNIRS analyses indicated that participants' gender and the form of social interaction that took place during the data acquisition could be related to the reported effect sizes. That is, a subgroup of analyses that examined male-female dyads found a larger overall effect than other subgroups of analyses that examined same-gender dyads, or that did not control participant gender. Similarly, a subgroup of analyses that examined inter-brain coupling during an interaction of limited medium (e.g., participants were asked not to talk during a face-to-face interaction), found a larger overall effect size than subgroups of other analyses examining other forms of interaction. However, these subgroups which found a larger overall effect than other subgroups consisted of a small number of analyses extracted from a limited number of studies. This can mean that these specific analyses in these particular papers involving opposite-gender dyads or an interaction of limited medium incidentally found a larger effect than other analyses in other studies. Therefore, it should not be concluded solely from this meta-analytic review that there is a specific relationship between dyad gender or interaction medium and inter-brain coupling.

Provided that the common methodological variance examined in the current meta-analysis did not explain the heterogeneity in the published findings, we must consider other contributing factors to the observed variation in the reported effect sizes. A possible explanation is due to the difference in data that were analysed in each study. Given that the variability observed in the current meta-analysis was largely associated with the study level rather than the analysis level, it is possible that the heterogeneity of the reported effect sizes from different studies is likely to be due to certain factors that exist differently between studies, such as the data cleaning process and the resulting data quality.

A particular note should be made on the impact of a unique combination of analytic decisions on the reported results in the literature. It has been shown that, in the fields of psychology and neuroscience, researchers have a great degree of freedom when choosing research methods to use (Botvinik-Nezer et al., 2020; Cohen, 2017; Gemignani & Gervain, 2021; Silberzahn et al., 2018; Stets & Reid, 2011). Importantly, previous studies where multiple teams of researchers analysed the same data found that each decision researchers make at each step of the analysis could alter the results of the analysis (Botvinik-Nezer et al., 2020; Gemignani & Gervain, 2021; Silberzahn et al., 2018). Botvinik-Nezer and colleagues (2020) asked 70 teams of researchers to analyse the same fMRI data. Comparing analytic pipelines and the results submitted, the authors reported that no two teams had used the identical analytic procedure, suggesting that there were substantial degrees of freedom on research teams in terms of analytic decisions. Further investigating potential moderators of the results, the authors found that the results of the analysis were biased by spatial smoothness, the software package used, and multiple test correlation methods. The effect of optimist bias by researchers was also identified, in such a way that a prior expectation for the hypothesis affected the results (Botvinik-Nezer et al., 2020). This indicates that differences in the results could arise from a stage of data handling (e.g., spatial smoothness), as well as researchers' bias, which is typically not clearly seen in the published manuscript. Other studies have shown that the process of data preprocessing could influence the analysis results. A recent study by Gemignani and Gervain (2021) focused on the steps of fNIRS data preprocessing to assess the effect of methodological choices on the results of analysis. They reviewed published studies in the last five years and extracted different methods to preprocess infant fNIRS data from 75 studies. The authors identified five pipelines that were used in 86% of the studies, whilst each of the pipelines included parameters that needed deciding by analysers. Moreover, applying each pipeline to synthetic and actual data, the authors found that each analysis method modified different characteristics of the data, leading to differences in the results of the analysis (Gemignani & Gervain, 2021). The preprocessing methods can vary in not only fNIRS and fMRI data but also EEG data, as manual artefact rejection is often recommended in addition to the use of automated rejection (Cohen, 2017). It has also been shown that differences in data inclusion or exclusion criteria and the resulting number of usable EEG data segments could modify the direction of the results in an EEG study with infants (Stets & Reid, 2011).

Importantly, such degrees of freedom regarding researchers' choice of methodology might be inevitable, and not to produce a certain type of evidence. Similar to a study done by Botvinik-Nezer et al. (2020), Silberzahn and colleagues (2018) asked 29 research teams involving 61 individuals to analyse the same data to address the same research question. The authors reported a large heterogeneity in analytic pipelines, whilst no significant methodological factors were found to explain the variability of the results (Silberzahn et al., 2018). In this study, the details of each analysis were shared among and reviewed by researchers from different research groups, ensuring that researchers made justifiable analytic decisions. Yet, these decisions could not be free from subjectivity, and contributed to different results. Therefore, it was concluded that some level of variations in the results of analyses on complex data might be unavoidable (Silberzahn et al., 2018). In sum, even using the same data collected in the same acquisition settings, the analytic procedures can differ, which can often lead to different conclusions. The current study reviewed studies that analysed different sets of data, each of which was collected using different setups for different research questions and hypotheses. When different elements of data and analyses are unique to each study, a considerable level of variability in the reported results, as found in the current meta-analysis, may not be surprising.

The substantial methodological variability and subsequent heterogeneity in the reported results identified in this meta-analysis indicate several steps that could be taken to ensure the interpretability of the results from hyperscanning studies, and therefore advance our understanding on human social brains that support complex and reciprocal human social interactions. First, establishing a standardised procedure for each step of research would be important. There has been a proposal for an analytic pipeline for hyperscanning research that could be applicable to other studies (e.g., Ayrolles et al., 2021; Barraza et al., 2019). It is hoped that future studies could use these shared pipelines and report their usability. A standardised protocol would improve the comparability of different studies, facilitating a reliable comparison across different results. Second, it is important to improve the transparency of our research practice, to ensure the reproducibility and replicability of research. Sharing the data and analysis code would enable other researchers to use other people's code on their own data, or run their own analysis code on the data of others. These would lead to more credible evidence by ensuring the results found do not result from methodological artefacts. Third, where possible, the data should be analysed by several multiple pipelines and ideally by more than one research team. Such a "multiverse analysis" has been proposed with differing methods, and would be of great use in hyperscanning research (Patel, Burford, & Ioannidis, 2015; Simonsohn, Simmons, & Nelson, 2015; Steegen, Tuerlinckx, Gelman, & Vanpaemel, 2016). In doing any of the above, we argue that transparent practice is paramount, and each decision during the process of research should be made according to their research aim and questions. It might sometimes be necessary to add modifications to the procedure commonly used in the past, or a standardised procedure. In such a case, the adaptations made and the justification for these changes should be clearly communicated, as each adaptation made might be what makes their research findings different from results from other studies.

Before concluding the current review, we must discuss some limitations of our analysis and its interpretation. First and foremost, this review should not be taken as an analysis to provide an accurate estimate of the inter-brain coupling between interacting individuals in social situations. As shown in the funnel plots, there is a publication bias in the estimates of overall effect sizes found in the current meta-analysis. Moreover, the current meta-analysis did not account for the direction of effects (i.e., all the effect sizes were extracted as an absolute value). Hence, we cannot conclude which condition is likely to find a stronger inter-brain coupling than others from this meta-analysis, further limiting the interpretability of the overall effect sizes provided by this meta-analysis. Furthermore, among the published results, the current meta-analysis only included significant main effects, excluding non-published results, non-significant effects, as well as significant interaction effects. This further limits the interpretability of the overall effect sizes provided in the current meta-analysis. This is particularly the case, because some researchers include, for instance, the brain regions of interest as a factor in their analysis, expecting to observe interaction to examine their hypothesis. Moreover, we rounded effect sizes in some of the included analyses, due to its reporting style (i.e., when the p-value was reported as $p <.05$, it was extracted as $p =.05$). This might have influenced the estimated overall effect sizes calculated in the current meta-analysis. These decisions were made according to our aim of assessing the effect of methodological variations on the heterogeneity in the reported results in the literature. Hence, we do not believe these limitations constrain the implications of this review. Nevertheless, a future meta-analysis including both published and unpublished studies and considering the effect direction would be beneficial for the field to assess the robustness of inter-brain coupling emerging on specific social contexts. Lastly, there was still likely to be a large heterogeneity in each subgroup of analyses, which might attenuate the strength of relationship between each methodological factor and reported results. Yet, subgrouping of analyses in the current review was conducted based on previous scoping reviews (e.g., Czeszumski et al., 2020), and it was necessary to include an adequate number of analyses in each subgroup to conduct a reliable and interpretable sub-group analyses. Hence, we argue that certain degree of heterogeneity in each subgroup was inevitable.

Despite these limitations, the current meta-analytic review emphasises the importance of developing ways to overcome the challenges of variability across every step of the methods in hyperscanning research, and the resulting heterogeneity in the results. With better validations of methods and analyses, hyperscanning research will have a great potential to uncover complex and dynamic brain systems that support spontaneous and reciprocal human social interactions.

Acknowledgement

We thank Hannah Thomas, Laura Diprossimo, Yazhi Liu for acting as second reviewers at the screening process. We also thank John Barbarock for his advice on the systematic literature search. We would like to extend our thanks to Dr Jade Pickering and Marta Topor for answering questions when finalising our pre-registered protocol using the NIRO framework. We also thank Marie Skłodowska-Curie Actions Innovative Training Network (MOTION) for their research funding and support.

Author contribution

Sayaka Kidby: design, systematic search, screening, data extraction, analyses, writing, review of the manuscript. Kirsty Dunn: design, review of the manuscript. Vincent Reid: design, review of the manuscript.

Chapter 3

Infants encode information from others differently according to their gaze cue congruency

Chapter Introduction

Given that our meta-analysis presented in the previous chapter suggested the lack of standardised protocols of hyperscanning research, it is best to explore various paradigms that could inform optimal designs of research investigating neural and cognitive processes underlying infant-adult reciprocal interactions.

When transitioning to a more naturalistic research paradigm involving less constrained and more dynamic social interactions between an infant and an adult, it is particularly important to identify reliable neural markers that we can turn to when conducting research in a more dynamic, and noisy environment, as otherwise the inferences we could make from the analysis can be limiting (Smith et al., 2021). For this, screen-based paradigms can be effective. While it typically positions infants in an 'observer' position, it can identify specific neural correlates of infants' cognitive processing of social environmental information, and demonstrate how infant cognition and the surrounding social environment are intertwined.

In this chapter, we present a study where we explored how 9-month-old infants attend to and encode information from two adults who give them gaze cues to an object location with different levels of accuracy. We showed how different characteristics about other people modulate infants' cognitive processes, pointing to the importance of studying infant cognition *within* social contexts, rather than separated from it.

Abstract

Children are selective learners from the first few years of postnatal development. It has been shown that infants as young as 8 months of age preferentially learn from 'reliable' informants over 'unreliable' informants (Tummeltshammer et al., 2014). Evidence suggests that attentional bias plays a role in infants' differential behaviour towards informants with different reliability (e.g., Heyes, 2017b), but less is known about what cognitive processes, beyond attentional processes, might underlie this selective learning in infants. To better understand infants' selective learning from reliable others over unreliable ones, the current study utilised event-related potentials (ERPs) to investigate whether 9-month-old infants' neural activity indicates the differentiation of two informants who give information with different congruency levels. Infants aged 9-months (*N*=22) were presented with stimuli depicting two females: one gave a congruent gaze cue to a target object location 100% of the time ('reliable' informant), and the other only gave a congruent cue 25% of the time ('unreliable' informant). Analysing infants' neural responses to the stimuli, we found differences in the mean and peak amplitudes of the Negative Central ERP component (Nc) as well as the mean amplitude of Positive Slow Wave (PSW) across conditions. The results indicate that infants allocated more attention to, and better encoded information from a 'reliable' informant when contrasted to an 'unreliable' informant. This study extends our understanding forward on how young infants collect and process social information, as the current study offers neural evidence that infants' attentional process is modulated by informant reliability, which further affects the degree to which information provided by an informant is encoded.

Introduction

We acquire and develop a significant amount of knowledge through social interactions, by listening to or observing other people (Harris, 2012). Especially during the early years of postnatal development, learning from other people is predominant and important (Meadows, 2010; Tomasello, 2014). When we learn from others, it is critical to be able to differentiate who is reliable as an informant, because information given by others could be a deception (Clément, 2010). An ability to assess other people's reliability is not only important for learning, but is also a social skill that plays an essential role in relationship building and cultural transmission (Ozer & Zheng, 2019). Whilst trust could lead to exploitation when used unethically (Yip & Schweitzer, 2015), with reliable others, it can reduce conflict (Zaheer, McEvily, & Perrone, 1998), enhance forgiveness after transgressions (Molden & Finkel, 2010), and has been associated with positive perceptions of one's relationships (Luchies et al., 2013; Rempel, Holmes, & Zanna, 1985). Hence, an ability to distinguish reliable others from unreliable ones is important beyond infancy.

The developmental trajectory of this skill has gained much attention for the last few decades. A growing body of studies has demonstrated that even pre-school children have the ability to preferentially learn from informants considered to be more 'reliable' than others (Harris, Koenig, Corriveau, & Jaswal, 2018; Mills, 2013). For instance, 3- and 4-year-olds only learn new object labels from a person who has previously labelled objects correctly all the time, but not from a person who has consistently labelled objects incorrectly (Koenig, Clément, & Harris, 2004). Furthermore, children aged 5- years are able to update their understanding about other people's reliability, not only prospectively but also retrospectively (Schütte, Mani, & Behne, 2020). Schütte and colleagues (2020) presented 5-year-old children with two informants who labelled a novel object differently. Following this, children saw each informant label a familiar object (e.g., apple) correctly or incorrectly, as an opportunity for them to determine which informant was trustworthy. Subsequently, it was tested which novel object label children learned, and children selectively learned the object–label pair provided by the speaker who turned out to be reliable. Before this study, it was proposed that children's selective learning is largely dependent on attentional bias, rather than their understanding of reliability (Heyes, 2017a, 2017b). Yet, this study demonstrates that it is linked to their higher socio-cognitive reasoning about other people's trustworthiness. In this study, reliability information was provided after the novel and contradicting information. Therefore, at the stage where children were faced with contradicting information, children should not know which informant was more reliable, and attention bias towards a reliable informant should not be present. Thus, children's selective learning is, at least in part, likely to be based on their meta-cognitive strategies such as trait reasoning.

Evidence also suggests that pre-verbal infants from 12 months of age might already be selective learners. Yet, what cognitive processes, such as trait reasoning, underlie infants' selective learning is less clear. For example, toddlers preferentially imitate a confident or a competent action agent, over modelling from an unconfident or incompetent agent (e.g. Buttelmann, Zmyj, Daum, & Carpenter, 2013; Stenberg, 2009; Zmyj, Buttelmann, Carpenter, & Daum, 2010). In a study by Zmyj et al., (2010), 14-month-old infants saw two adults. A 'competent' adult manipulated a familiar object (e.g., shoes) with a confident facial and vocal expression and successfully achieved an action goal using the object (e.g., put on shoes), while an 'incompetent' adult acted on the same object with a facial and vocal expression showing uncertainty and failed to achieve an action goal. They found that infants were more likely to imitate a competent adult as compared to an incompetent adult. Yet, it is possible that children in the study simply understood the goal of action and tried to achieve it, rather than selectively imitated a more competent person (Heyes, 2017b). Therefore, the interpretation of this result in the context of infants' understanding of other people's reliability needs to be taken with caution. Yet, this study does suggest that children as young as 14 months of age do not simply imitate everything they observe.

In a similar vein, it has been reported that 12-month-old infants looked more towards an expert than a non-expert when they were presented with an ambiguous toy (Stenberg, 2013). This looking behaviour under an uncertain situation is typically called *social referencing* (Rochat, 2014; Walden, Kim, McCoy, & Karrass, 2007; Walden & Ogan, 1988). It has been demonstrated that infants engage in this looking behaviour when they expect to solicit information from others (Feinman, Roberts, Hsieh, Sawyer, & Swanson, 1992). Hence, infants' selective social referencing might indicate their understanding of who are more 'reliable' as an informant, or who are more likely to be able to give information. A more recent study by Bazhydai and colleagues (2020) provided supporting evidence. In their study, 12-month-old infants were introduced with two experimenters and a novel object. One of them played a role of a 'reliable' informant, and labelled the novel object. The other played a role of an 'unreliable' informant, and showed ambiguous expressions about the object without labelling. After infants saw these two experimenters labelling or not labelling the object, infants were presented with two objects, one from the previous session ("target object") and the other being novel. They were then asked to choose the target object by their parent. The results showed that infants reliably looked towards the 'reliable' informant first before the 'unreliable' informant. Furthermore, the total duration of looks towards the 'reliable' informant was longer than the total look duration towards the 'unreliable' informant (Bazhydai, Westermann, & Parise, 2020). Perhaps though, it is possible that infants' differential visual engagement towards the two experimenters in this study might reflect infants' associative learning and attentional biases rather than their understanding of psychological concepts such as reliability. In this study by Bazhydai et al., (2020), the 'reliable' experimenter repeated the same object label in their utterance (e.g., "Look, this is a ball!"; "Wow, a ball!"; "What a nice ball!"; "I like this ball!", "Where is the ball?", "A ball!"), whereas the 'unreliable' experimenter made different comments each time she spoke (e.g., "Look at this!"; "Oh Wow!"; "What is this?"; "Hmm..."; "This is nice!"; "I like this!"). Infants are known to be sensitive to statistical regularities, and they pay more attention to predictable information sources (Heyes, 2017b). Therefore, whilst this study demonstrates that infants show differential behaviour to those who consistently have information to offer and to those who do not, the result might simply reflect infants' understanding of which experimenter can give more predictable responses than the other.

Research has suggested that even younger infants might also be able to tell 'reliable' informants from 'unreliable' ones (Tummeltshammer, Wu, Sobel, & Kirkham, 2014). Critically though, young infants' ability to discriminate between 'reliable' and 'unreliable' others has been discussed in relation to their sensitivity to statistical regularity, rather than their cognitive skills of trait reasoning (Haith, 1993; Kirkham, Slemmer, & Johnson, 2002; Tummeltshammer et al., 2014). Tummeltshammer and colleagues (2014) familiarised 8-month-old infants with two adults using screen-based stimuli. There were four boxes positioned at each corner of the screen. One of the adults always looked at the box where an animal image appeared shortly after she looked there ('reliable' informant), whilst the other adult only looked at the box of the animal 25% of the time, and looked away from the corner 75% of the time ('unreliable' informant). After this familiarisation, they assessed infants' gaze following behaviour towards these two types of informants. They found that infants only followed the gaze of the 'reliable' informant, but not the gaze of the ‘unreliable’ informant (Tummeltshammer et al., 2014). This study illustrates that infants as young as 8 months of age can discriminate a ‘reliable’ informant from an ‘unreliable’ informant based on the history of gaze cue congruency, and modify their gaze following behaviour according to the ‘reliability’ of informants.

Whilst infants’ sensitivity to statistical regularity and attentional bias seems to play a role in infants’ selective behaviour, it has also been suggested that evaluating other people’s reliability may well be more than mere associative learning. Tummeltshammer and colleagues (2014) also investigated whether 8-month-old infants learn to follow cues given by inanimate arrow-like symbols. The results suggested infants’ gaze indeed follow “reliable” arrows but not “unreliable” arrows. What was different between eye gaze and arrow cues was the generalisation of gaze following. In their studies, during the familiarisation phase, reliable and unreliable informants, both adults and arrows, cued one of the two specific boxes, and they never cued the other two boxes. During a test trial, this was continued, and informants only cued one of the two boxes which they had cued during the familiarisation trials. During a generalisation test phase though, informants could cue any of the four boxes, not limited to the two boxes that they had previously cued during the familiarisation and test trials. During the test trial, infants followed both eye gaze and arrow cues. During the generalisation trial, however, the infants followed eye gaze cues but did not follow arrow cues. This means that infants are able to form associations between certain boxes and the direction of the arrows, but the function of an arrow as a cue to a target location for information delivery cannot be generalised. On the other hand, infants understand the cueing function of human eye gaze and were able to generalise the function so it can be applied to a novel situation. However, gaze and arrow cues differ in terms of familiarity, not only whether they convey social information (Heyes, 2017b; Tummeltshammer et al., 2014). Hence, it Is possible that infants were only able to generalise the cueing function of the gaze because they were already familiar with adults face directing their gaze towards locations and objects, whereas they did not have enough exposure to such a novel cue as an arrow-like shape used in the study for them to generalise its function (Elsner & Pauen, 2007; Tummeltshammer et al., 2014). Therefore, it is difficult to argue the difference in the generalisability is solely due to the difference between social or non-social agents, and hence, due to infants’ understanding of social and psychological traits, such as reliability.

In sum, the existing studies suggest that infants as young as 8 months of age can discriminate ‘reliable’ informants from ‘unreliable’ informants, at least in part, based on attentional bias towards consistent and predictable information sources, as well as familiarity (Heyes, 2017b; Tummeltshammer et al., 2014). However, an issue that remains unanswered is whether, and how, infant cognition might differ when they are faced with reliable and unreliable others, beyond attentional processes.

To address the question and explore infants’ online cognitive processing while young infants track and process informant reliability, an electroencephalogram (EEG) can provide insights as it can be recorded concurrently with a stimuli presentation. Event-related potential (ERP) techniques have been commonly used to identify infants’ cognitive activity. An ERP represents an electrical brain activity that is time-locked to a specific aspect of a stimulus, such as its onset (Rugg & Coles, 1996).

Given the evidence that infants’ visual attention (i.e., looking behaviour) is modulated by informant reliability (Tummeltshammer et al., 2014), it is plausible that infants’ covert attentional processes which are reflected in a neural activity are also modulated by informant reliability. The ERP component which has typically been linked to attentional processes is Negative Central component (Nc), with a larger-amplitude Nc indicating more attentional resources allocated to process the presented stimuli when contrasted with a smaller-amplitude Nc for another condition (de Haan, 2013; Reynolds & Richards, 2005). In infants, it typically peaks between 400ms and 800ms following the stimulus onset, and is commonly observed to be maximal over fronto-central scalp regions (Snyder, Webb, & Nelson, 2002). Investigating the function of Nc in terms of infants’ attention allocation to salient stimuli, studies have linked a larger-amplitude Nc to a longer fixation of infants’ visual attention (Ackles & Cook, 1998, 2007; J. H. Karrer, Karrer, Bloom, Chaney, & Davis, 1998; Reynolds, Courage, & Richards, 2010). In studies using an oddball paradigm, where, typically, one of the two stimuli (‘standard’) is more frequently presented than the other (‘oddball’), both a larger-amplitude Nc and longer looking time have been found for the ‘oddball’ stimulus than the ‘standard’ (e.g., Courchesne et al., 1981; Hill-Karrer et al., 1998; Hunter & Karrer, 1993). Yet, studies using other paradigms have reported that the Nc amplitude and looking duration are not significantly correlated (de Haan, Belsky, Reid, Volein, & Johnson, 2004; Nikkel & Karrer, 2009), and that Nc amplitude can be different across conditions even infants’ looking behaviour does not differ (Parise et al., 2008). A possible explanation for the mixed results regarding the relationship between looking time and Nc amplitude can be that large individual differences in Nc amplitude overwhelm more subtle modulations in attentional processes that the Nc reflects, obscuring relations with looking time (de Haan, 2013). Reynolds and colleagues (2010) investigated the association between infants' Nc and attention by using heart rate as an index of attention. The greater Nc amplitude was found when infants were attentive than when in attentive, which indicates that the the Nc is likely to be functionally related to infants' attention allocation (Reynolds & Richards, 2005). Furthermore, the authors analysed the cortical source of the Nc component, which found its sources in inferior prefrontal areas, superior prefrontal areas, and the anterior cingulate. As these areas have been linked to a variety of cognitive functions, including executive attention, recognition memory, shifts of attention, suppression of saccades (Duncan & Owen, 2000), the finding of this cortical source analysis also supports the proposal that these areas are involved in sustained attention (Reynolds et al., 2010; Reynolds & Richards, 2005). Whilst further investigations would be necessary to exactly identify what specific aspect of attention the Nc or its amplitude reflects, the existing literature generally indicate that the Nc is a valid and informative neural measure of infant attention (Reynolds et al., 2010; Reynolds & Richards, 2005). In the context of eye gaze cueing paradigms with infants, the amplitude of the Nc has been found to differ between congruent and incongruent gaze. For example, Hoehl and colleagues (2008) presented 4-month-old infants with static images containing a face and an object displayed on either side of a face. The eyes in the face were either looking at the object, or looking away from the object. The amplitude of the Nc was found to be larger for the gaze averted from the object (incongruent gaze) than the gaze directed to the object (congruent gaze). This suggests that infants allocated more attention to incongruent gaze as compared to congruent gaze. In the context of selective learning based on the gaze cue congruency of informants, previous studies have reported that infants' visual attention is guided towards more 'reliable' informants who consistently give congruent gaze cues (Heyes, 2017b; Tummeltshammer et al., 2014). Therefore, it is possible that the Nc amplitude might be larger for congruent informants as compared to incongruent ones, reflecting infants' attentional bias towards 'reliable' informants.

Infants' attention is closely tied to information encoding and visual recognition memory, as encoded information, or the content of memory, can generally be inferred from visual attentional target (de Haan, 2013; Reynolds et al., 2010; Reynolds & Richards, 2005). Thus, it is plausible that infants' information encoding and memory are also modulated via the modulation of attentional processes according to informant reliability. The Positive Slow Wave (PSW) is an ERP component which typically occurs between 600ms and 1000ms after the stimulus onset and is most prominent over fronto-temporal channel regions (Reid et al., 2004), and has been linked to the degree of information encoding (de Haan, 2013; de Haan & Nelson, 1999; Nelson, 1994, 1997). Using a modified three-stimulus oddball paradigm, in which infants are first familiarised with 'standard' and 'oddball' stimuli and then presented a novel infrequent stimulus along with the standard and oddball ones, Nelson and Collins (1991, 1992) found no differences in Nc amplitude in 6-month-olds. Nevertheless, they found differential slow wave components across stimuli; infants showed PSW in response to the infrequent-familiar (familiar 'oddball') stimuli, whereas they showed a different ERP component (Negative Slow Wave; NSW) to the infrequent-novel stimuli. This replicates the results reported by Karrer and Ackles (1987) with 6- and 7.5-month-old infants. Thus, it has been proposed that NSW reflects infants' detection and processing of novelty, whereas PSW indexes an updating of working memory for a stimulus that had previously been partially processed (Reynolds & Richards, 2005). Importantly, differences in these cognitive processes indexed by Positive and Negative Slow Wave components are not reflected by a difference in Nc amplitude, pointing to the importance of exploring these components separately to better understand different aspects of cognitive processes at play. The sources of PSW is thought to involve temporal cortical and subcortical areas, including the hippocampus (Nelson, 1994; Reynolds & Richards, 2005). Given that adult studies have reported that the hippocampal-parahippocampal region is involved in encoding of social information such as faces (Grady et al., 1994), this is consistent with the proposal that PSW is functionally linked to information encoding which further informs recognition memory establishment (Reynolds & Richards, 2005).

These adult studies linking PSW to social information processing point to the potential use of PSW as a sensitive index of social information processing. In studies with infants, PSW has been associated with the processing of congruent and incongruent gaze cue. Reid and colleagues (2004) presented 4-month-old infants with videos of an adult face whose eye gaze moving either towards or away from the nearby object. After the face disappeared, the object was centrally displayed and EEG was recorded while the infant attended to the object. They found the enhanced amplitude of PSW for objects which was not cued by the adult, compared to objects which was the adult's attentional target. Given that the amplitude of PSW becomes larger for an object which has partially been encoded, this result suggests that cued objects are better encoded than uncued objects, which are consistent with the literature stressing the role of other people's gaze that biases infants' attention (e.g., Becchio, Bertone, & Castiello, 2008; Reid et al., 2004; Striano et al., 2006). While the examination of PSW does not directly clarify whether young infants have an understanding of reliability as adults and older children would do (e.g., Schütte, Mani, & Behne, 2020), it is possible that differential PSW amplitudes can indicate differences in infants' cognitive processes, such as information encoding and memory, which may well inform infants selective behaviour towards different types of informants, beyond attentional bias.

The current study aims to identify the neural correlates of young infants' processing of information given by 'reliable' and 'unreliable' informants. More specifically, this study explores how the brain activity of 9-month-old infants might differ in response to different levels of congruency of other people's gaze cues, based on the gaze cueing paradigm proposed by Tummeltshammer et al. (2014). If 9-month-old infants can discriminate between congruent (reliable) and incongruent (unreliable) informants, their neural responses to these informants should dissociate from each other. Based on the existing studies utilising ERPs that identified the neural correlates of infants' discrimination between congruent gaze (gaze towards the object) and incongruent gaze (gaze averted from the objects), we hypothesised that the larger amplitude of the Nc would be observed for the congruent informant compared to the incongruent informant, reflecting the different level of attention allocation based on the gaze cue congruency in such a way that infants' attention is guided towards a congruent informant when contrasted with an incongruent informant (Tummeltshammer et al., 2014). We also expected to find an enhanced PSW amplitude for the incongruent informant compared to the congruent informant, which would be indicative of different degrees of information encoding according to the cue congruency.

Methods

Participants

The final sample consisted of 22 infants aged 9 months (11 females), with the average age being 271 days (ranging from 254 days to 288 days, *SD* = 10.24 days). All infants were born full term (37-41 weeks) and in the normal range for birthweight (over 2500 grams). The age group was selected based on a study by Tummeltshammer et al. (2014) which showed that the ability to track informant's reliability is present at 8 months of age. All infants were recruited through phone calls and emails from a database consisting of parents who had expressed an interest in taking part in psychological research. The participating families were predominantly white and middle class. An additional 31 infants (10 females) participated but were excluded from the final sample due to fussiness (*n*=1), an inadequate number of trials (i.e., less than five usable trials) for data averaging as a standard ERP analysis procedure (*n*=27), or a technical or experimental error (*n*=3). The attrition rate of 58.4% is within a typical range of attrition rate for an infant EEG study (40-65%) (e.g. Hoehl et al, 2014; Reid et al., 2004).

Prior to recruitment, ethical approval regarding the study procedure (e.g., recruitment, research methods and data handling throughout the study) was sought and obtained from the Lancaster University Faculty of Science and Technology Research Ethics Committee. All the parents whose child participated in the current study were given a full description about the study from the experimenter, and gave a written informed consent on behalf of their child.

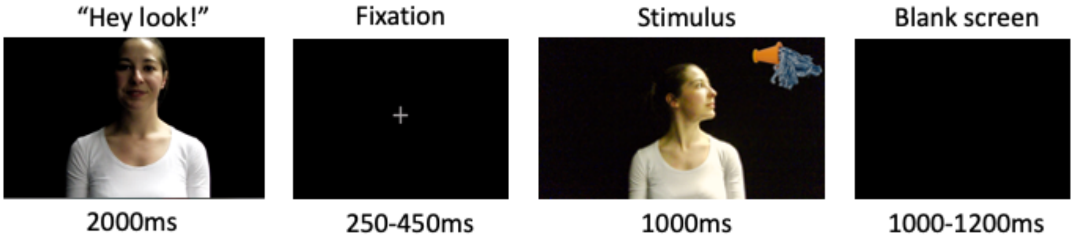
Stimuli and apparatus

The stimuli were created at Lancaster University, modelling the stimuli reported in Tummeltshammer et al. (2014) but with modifications. The stimuli comprised 32 static images in total. Each image displayed one female looking at one of the four corners of the screen as well as one object appearing at one of the screen corners. In 16 images, the same female looked at the object appearing on one of the corners 100% of the time (i.e., congruent informant). In the other 16 images, a different female looked at the object 25 % of the time and looked away from the object 75% of the time (i.e., incongruent informant). The order of the presentation was determined using pseudo-randomisation with no constraints on randomisation procedure. The angle between the centre of the object and the eyeline of the female was approximately 11 degrees.

Before the experimental stimuli, a 2-second video clip was presented, in which a female said “hey, look!” with her eyes directed to the infant. This was to draw infants' attention and signal them that the female was about to give them information. Following the short video clip, the fixation cross was presented, of which duration was randomly assigned between 250ms and 450ms. After the stimuli image, a black blank screen was presented for a duration randomly assigned between 1000 and 1200ms, before the next video clips commenced.

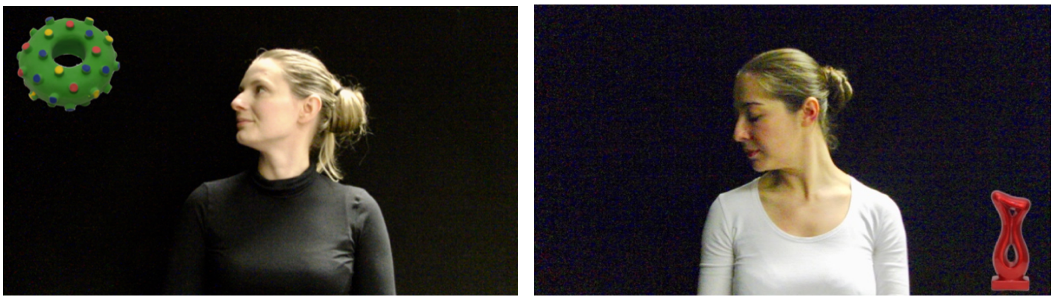
Each image subtended a visual angle of approximately 13° horizontally and 15.5° degrees vertically on a 22-inch screen (resolution 1350 x 1080 pixels). The distance between the monitor and the infant was approximately 70cm.

Figure 3-1. Illustration of the trial sequence.



Note. A two-second video clip was played in which a female said "hey, look!" with her eyes directed towards the infant, following which the fixation cross was presented for a randomly assigned duration between 250ms and 450ms. The stimulus showing a congruent or incongruent informant appeared after the fixation cross for 1000ms. Between trials, a blank screen was presented for a randomly assigned duration between 1000ms and 1200ms.

Figure 3-2. Examples of stimuli for the congruent informant (left) and the incongruent informant (right).



Procedure

In an experimental room, infants sat on their parent's or caregiver's lap in front of a monitor, and the EEG cap was put on the infant. Infants were assigned to one of the two groups according to which female was a congruent informant. The assignment was to cancel out the potential confound of a specific individual being a congruent (or an incongruent) informant across participants. The number of infants in each group was nearly equal when recruitment and data acquisition ceased (26 infants in one group and

27 in the other). All the infants watched stimuli of both types of informants, and contributed data to both conditions.

In order to minimise possible distractions for infants, lighting conditions in the experimental room were kept low, and a partition wall was placed to separate the experimenter and EEG equipment from the infant. The experimenter sat behind the partition wall out of view of the infant, and monitored the infant's behaviour and gaze direction through a camera placed beneath the screen that presented the stimuli. When the infant did not seem to look at the stimuli, the experimenter used an attention-grabbing video clip (e.g., a spiral moving to music) to orient the infant's attention back to the stimuli. The experimenter also took a short break when she considered it to be effective to draw the infant's attention back to the stimuli. Infant behaviour was video-recorded throughout the data acquisition session using the same camcorder. The video recordings were used to code infants' visual attention towards the stimuli on a trial-by-trial basis offline after the infant had left the laboratory.

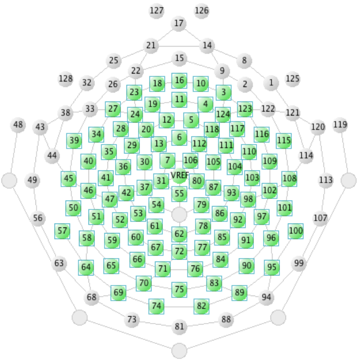
Parents were instructed not to let their child move around to an excessive level. They were also asked to avoid interacting with the child during the stimuli presentation. Infants who were included in the final analyses (*N*=22) saw 82 images on average (ranging from 66 to 110, *SD* = 12.84) before infants were too distracted to reliably maintain their attention on the screen. The total number of trials infants attended to did not differ across conditions (*t*(21) = 0.87, *p* = .40).

EEG recordings and analysis

Infants' electrical brain activity was measured using a 128-channel Geodesic Sensor Net while infants watched the video stimuli. EEG data were amplified with an EGI Net Amps 400 amplifier with a sampling rate of 500Hz (HCGSN 130, EGI, Eugene, OR, USA). The four electrooculogram (EOG) channels (channels 125, 126, 127, 128) were disconnected and not used in the current study to avoid excessive discomfort for infant participants, having electrodes close to their eyes. Therefore, the data for the current study were collected by 124 channels in total. Furthermore, the electrodes from the most anterior and posterior areas were also excluded from the final analysis due to high noise caused by poor contact with the scalp, which is a common data management process in infant EEG studies (e.g., Bakker, Sommerville, & Gredebäck, 2016; Gredebäck et al., 2015;

Gredebäck, Melinder, & Daum, 2010). Thus, 38 electrodes in total were excluded from the final analyses. The final analyses were conducted over the data collected from 90 channels (see Table 3-3).

Figure 3-3. Channels included in the final examination.



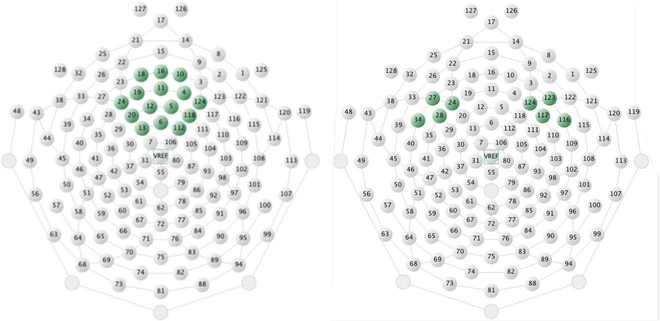
EEG was referenced online to the vertex electrode (Cz), with an analogue bandpass filter being applied (0.1 to 100 Hz) when storing the data for subsequent offline analysis. Signals (triggers) were sent to the EGI system simultaneously when the stimulus image was presented on the monito to segment the data into epochs based on the stimuli onset. The data were processed prior to inspections for ERPs using NetStation 4.5.4 Waveform Tools. The data were first filtered using 0.3Hz – 60Hz bandpass, which then were segmented so each segment would include from 200ms before the stimuli onset to 1000ms after the stimuli onset. This segmentation was based on the literature which reported that components of interest in the current study (Nc, PSW) were observed within 1000ms after the stimuli onset (Kaufman et al., 2005; Parise et al., 2008; Reid et al., 2004; Senju et al., 2006; Striano et al., 2006). After the segmentation, the automatic artefact detection tool was applied to each segmented EEG data (i.e., trial), which marked channels in each trial as “bad” when the amplitude exceeded the set threshold of +/- 200 microvolts for the entire segment, with a moving window average of 80ms. This automatic artefact detection was manually and visually inspected afterwards to improve the inspection by pre-programmed algorithms, because these are typically developed for adult EEG data, and are not always sensitive to artefact in infant EEG (for further information on infant ERP methods, see Hoehl & Wahl, 2012).

Based on the manual artefact detection, trials which included more than 12 channels (i.e., more than 10% of the 124 channels used for data collection) containing excessive level of noise were excluded for the following steps of the analysis. At this time point, data from 31 participants out of 53 participants who contributed less than five usable trials per condition were excluded from the subsequent analyses. This data inclusion criteria of five usable trials per condition was determined based on previous studies which showed that reliable and interpretable analyses could be conducted from three to seven usable trials per condition in an infant EEG study (Kaduk et al., 2016; Kaduk, Elsner, & Reid, 2013; Stets & Reid, 2011).

After selecting the participants whose data were considered relatively low in noise (i.e., less than 10% of the total channels marked as bad), the channels marked as bad were interpolated by using the average of surrounding electrodes. Following the interpolation, the channels were re-referenced to the average electrode. This was followed by the baseline correction. The baseline used for the current study was 200ms before the stimulus onset, following standard ERP procedures (Luck, 2005). Then an individual average for each participant was calculated. A grand average was finally computed using the 22 individual averages.

Based on the literature, the Negative Central component (Nc) was examined over fronto-central electrode sites (see Figure 3-4). The time window from 400ms to 600ms after stimulus onset was chosen for the examination of the Nc modelling Hoehl et al. (2008). The Positive Slow Wave (PSW) was assessed over fronto-temporal electrode sites (see Figure 3-4), using the mean amplitude over a time window from 600ms to 1000ms based on the previous literature (Hoehl et al., 2008; Nelson & de Haan, 1996; Webb, Long, & Nelson, 2005).

Figure 3-4. Channels inspected and used for statistical tests.



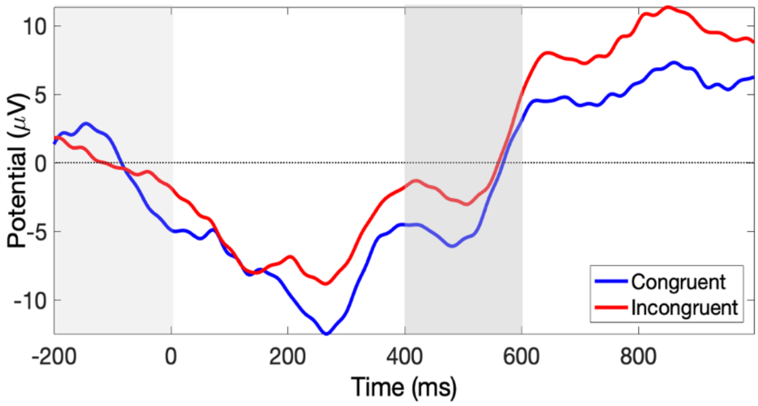
Note. The left plot shows the fronto-central channel regions used to examine the Nc occurring between 400 and 600ms after the stimuli onset. The right plot shows the fronto-temporal channel regions used to inspect the PSW occurring between 600ms and 1000ms after the stimuli onset.

Results

The Negative Central component (the Nc)

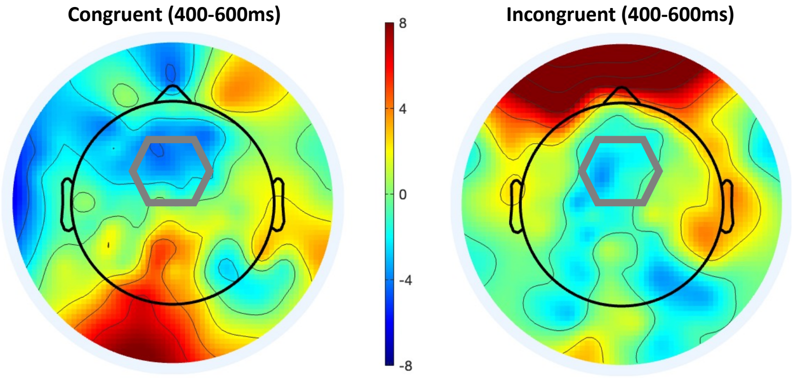
Examining the mean amplitude of the Nc across the time window of 400ms to 600ms after stimulus onset over fronto-central electrode sites, a paired-sample *t*-test found a significant difference in the mean amplitude between the congruent ($M_{congruent}$ = 7.03 [μV], $SE_{congruent}$ = 1.00) and the incongruent condition ($M_{incongruent}$ = 10.30 [μV], $SE_{incongruent}$ = 1.27) (*t*(21) = -2.25, *p* = .035, Cohen's *d* = - 0.48), indicating that the amplitude of the Nc was enhanced during the congruent condition compared to the incongruent condition. The peak amplitude (i.e., minimum amplitude) within the selected time window was also examined, because the use of the peak amplitude, not the mean, is also a common procedure in statistically assessing differences in ERP waveforms. A paired-sample *t*-test found a significant difference in the peak (minimum) amplitude between the congruent ($M_{congruent}$ = 1.72 [μV], $SE_{congruent}$ = 1.04) and the incongruent condition ($M_{incongruent}$ = 5.58 [μV], $SE_{incongruent}$ = 1.50) (*t*(21) = -2.33, *p* = .030, Cohen's *d* = -0.50). This is consistent with the result of the comparison using the mean amplitude, and demonstrates that the amplitude of the Nc was enhanced during the congruent condition compared to the incongruent condition.

Figure 3-5. Grand-average waveforms on the congruent (blue) and incongruent (red) trials over fronto-central electrode sites.



Note. The darker grey shading indicates the time window used to examine the Nc (400ms to 600ms), whereas the lighter grey shading marks the baseline period (-200ms to 0ms). The time = 0 refers to the time at which the static image in each trial appeared on the screen.

Figure 3-6. The average scalp distribution of the grand average waveform, within a timeframe between 400ms and 600ms after stimuli onset, for the congruent (left) and incongruent (right) conditions.

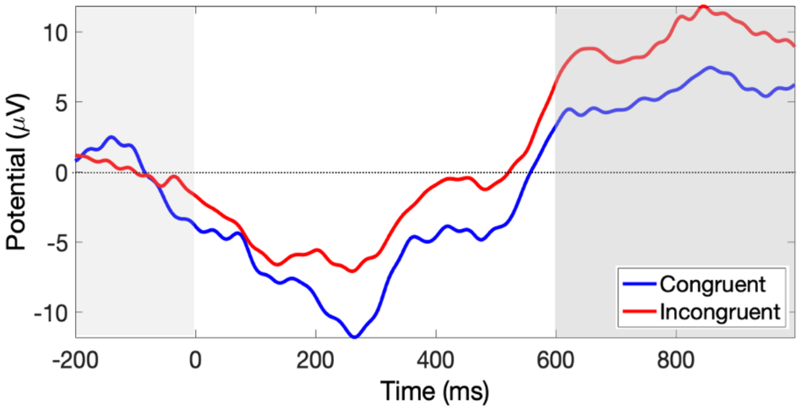


Note. The grey area represents the fronto-central electrode regions examined.

The Positive Slow Wave (PSW)

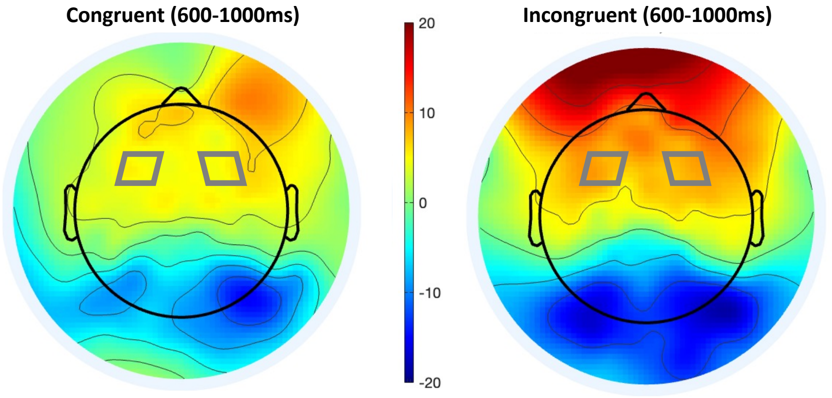
Regarding the mean amplitude of PSW across the time window between 600ms to 1000ms after stimulus onset over fronto-temporal electrode sites, a paired-sample *t*-test comparing the congruent and incongruent conditions found a significant difference between the conditions, indicating that there was a larger PSW effect during the incongruent condition ($M_{incongruent}$ = 9.49[μV], $SE_{incongruent}$ = 1.13) than the congruent condition ($M_{congruent}$ = 6.17 [μV], $SE_{congruent}$ = 1.13) ($t(21) = -2.53$, $p = .020$, Cohen's $d = -0.54$).

Figure 3-7. Grand-average waveforms on the congruent (blue) and incongruent (red) trials over fronto-temporal electrode sites.



Note. The darker grey shading indicates the time window used to examine the positive slow wave (600ms to 1000ms), whereas the lighter grey shading marks the baseline period (-200ms to 0ms). The time = 0 refers to the time at which the static image in each trial appeared on the screen.

Figure 3-8. The average scalp distribution of the grand average waveform, within a time frame between 600ms and 1000ms, for the congruent (left) and incongruent (right) conditions.



Note. The grey area represents the fronto-central electrode regions examined.

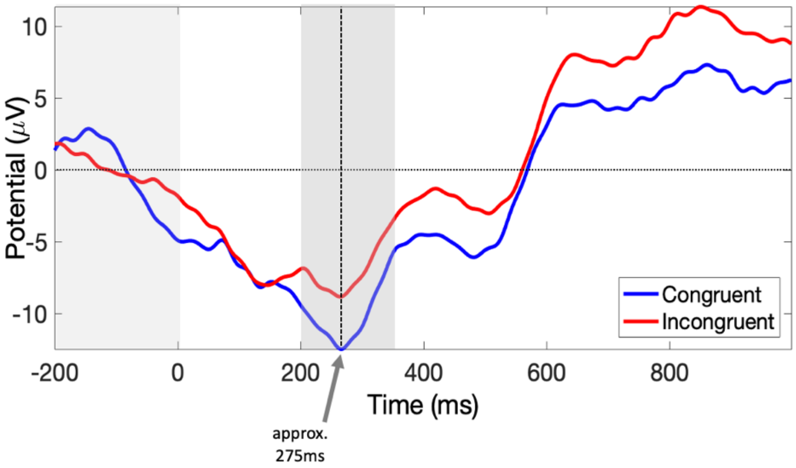
Early-latency negative deflection

Visually inspecting the grand average waveforms, the negative deflection occurring before 400ms was prominent. The morphology of this waveform is difficult to interpret as this was at an earlier latency than the conventional Nc. Other visually derived ERPs with infants have induced early components, such as the Pb, a positive deflection occurring approximately 200 to 400 milliseconds after stimulus onset (Webb et al., 2005). Nevertheless, these are reported in so few published papers that it is still difficult to label this component (Webb et al., 2005). We therefore conducted an exploratory analysis examining the difference in the mean amplitude of this negative deflection, within the time window from 200ms to 350ms. The time window was chosen based on the visual examination of the grand average waveforms, and the deflection was examined over the fronto-central electrode sites which were identical to the regions used for the examination of the Nc.

A paired-sample *t*-test comparing between the congruent and incongruent conditions indicated that the mean amplitude of the negative deflection peaking at approximately 275ms after stimulus onset was larger in the congruent condition ($M_{congruent}$ = -10.18[μV], $SE_{congruent}$ = 0.97) than the incongruent condition ($M_{incongruent}$ = -7.11[μV], $SE_{incongruent}$ = 3.82) ($t(21) = -2.86$, $p = .009$, Cohen's $d = -0.61$). The peak (minimum) amplitude within the specified time window was also examined. A paired-sample *t*-test found a significant difference in the peak (minimum) amplitude between the congruent ($M_{congruent}$ = -14.85 [μV], $SE_{congruent}$ = 0.97) and the incongruent condition ($M_{incongruent}$ = -12.14 [μV], $SE_{incongruent}$

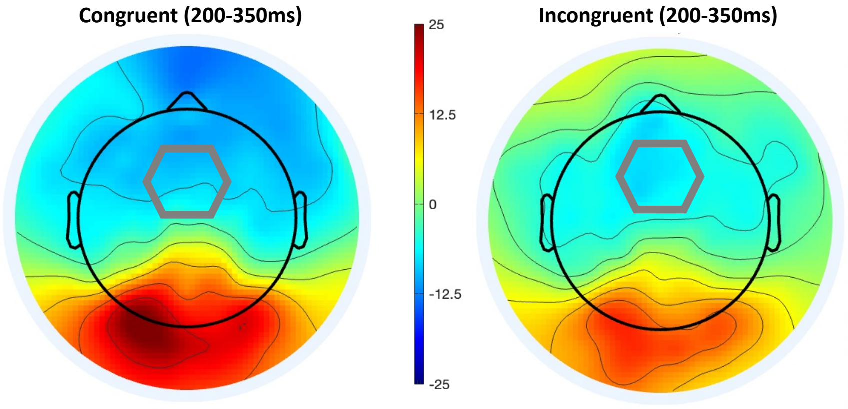
= .82) (*t*(21) = -2.62, *p* = .016, Cohen's *d* = -0.56). This is consistent with the result of the comparison of the mean amplitude, demonstrating that the amplitude of this early-latency negative deflection was increased during the incongruent condition compared to the congruent condition.

Figure 3-9. Grand-average waveforms on the congruent (blue) and incongruent (red) trials over fronto-central electrode sites.



Note. The darker grey shading indicates the time window used to examine the early-latency negative deflection (200ms to 350ms), whereas the lighter grey shading marks the baseline period (-200ms to 0ms). The time = 0 refers to the time at which the static image in each trial appeared on the screen.

Figure 3-10. The average scalp distribution of the grand average waveform, within a time frame between 200ms and 350ms, for the congruent (left) and incongruent (right) conditions.



Note. The grey area represents the fronto-central electrode regions examined.

Discussion

Previous studies have shown that young children and even preverbal infants are selective learners (e.g., Mills, 2013). During infancy when children are still developing their language and social skills, they might make use of statistical regularity to navigate their attention towards a 'reliable' informant, who consistently gives congruent cues to information (Tummeltshammer et al., 2014). It has also been shown that infants' sensitivity to other people's gaze cue congruency can modulate infants' gaze following behaviour (Tummeltshammer et al., 2014). Yet, less is known about what cognitive processes beyond attentional processes underlie infants' discrimination of 'reliable' and 'unreliable' informants. Utilising electrophysiological measures, the present study aimed to explore the attentional and cognitive processes of how infants track and encode people's reliability as an informant, which was reflected in their gaze cue congruency. Our analysis of event-related potential components (ERPs) found a difference in the Nc and PSW amplitudes between the processing of the congruent and incongruent informants, suggesting differential attentional and information encoding processes in infants depending on the gaze congruency of the informants. Moreover, there was a significant difference in neural activity occurring at an earlier latency than the Nc. In the following, we discuss these three neural components that indicate infants' differential processing of congruent and incongruent informants.

First, infant brains exhibited an enhanced amplitude of the Nc while infants were observing the congruent informant, compared to the incongruent informant. Previous studies have associated infant Nc with their attentional allocation (e.g., Striano et al., 2006; Parise et al., 2008). Therefore, consistent with previous research by Tummeltshammer et al. (2014), the current study showed that infants' attention allocation was guided towards a congruent informant more when contrasted with an incongruent informant. This might seem contradictory compared to previous findings that the greater Nc was found in response to incongruent gaze than congruent gaze (Hoehl et al., 2008). However, a greater Nc amplitude has also been found for more socially relevant and salient stimuli (e.g. parent's face, familiar toy) than novel stimuli (e.g. stranger's face, novel toy) (de Haan & Nelson, 1997, 1999). Studies utilising looking measures also indicate that infants' attention is affected by various factors, such as age (Wetherford & Cohen, 1973), and the degree to which the information is encoded (Cashon & Cohen, 2000; Hunter & Ames, 1988; Schöner & Thelen, 2006). Hence, it is not always straightforward to compare different studies involving children of different age and different stimuli, in terms of which stimulus should attract a greater allocation of infants' attention. Furthermore, a different study has reported that a larger Nc amplitude was found while infants interacted with an adult who engaged in joint attention with an infant, compared to when they interacted with an adult who only looked at the monitor presenting stimuli but not at the infant (Striano et al., 2006). Joint attention involves direct eye contact with infants, which is considered to signal other people's communicative intention (Csibra & Gergely, 2009). Hence, infants might perceive an adult who engages in joint attentional interaction as more socially salient and important, as compared to an adult who does not make a direct eye contact. In the context of the current study, our finding of Nc amplitude difference may well indicate that infants perceive a congruent informant to be more socially relevant and salient, as compared to an incongruent informant, and hence allocated more attention to the congruent informant than the incongruent informant.

Second, there was a larger Positive Slow Wave effect (PSW) for the incongruent informant than the congruent informant. The PSW has been thought to reflect memory updating, and the mean amplitude of PSW has been found to be greater for stimuli which were previously encoded only partially (Reid et al., 2004; Snyder et al., 2002). In previous studies investigating infants' object processing using a gaze cueing paradigm, the amplitude of PSW is often found to be larger for an uncued object compared to a cued object (Hoehl, Wahl, Michel, & Striano, 2012; Kopp & Lindenberger, 2011; Reid et al., 2004). Therefore, the larger-amplitude PSW for the incongruent informant observed in the current study suggests that a consistent gaze cueing pattern of the congruent informant aids infants' information encoding, and the inconsistent accuracy of gaze cues from the incongruent informant might make it more difficult for infants to process. This finding has advanced our knowledge by demonstrating a difference in gaze congruency between informants modulates not only attentional processes but also the degree of infants' information encoding. This is consistent with the predictive processing perspective towards infants' learning, which argues that the infant brain is sensitive to statistical regularities and their learning occurs based on predictions (Köster, Kayhan, Langeloh, & Hoehl, 2020). As the gazing pattern of the congruent informant was consistent in such a way that they always looked at the object, infants' cognitive processing might be more attuned to the congruent informant when contrasted with the incongruent informant.

Lastly, we found differential neural activity at an earlier latency than the Nc, peaking at around 275ms after the stimuli onset. In terms of morphology, it was very similar to the Nc component, being a negative deflection maximal over fronto-central electrode regions. However, because of its early latency, it is difficult for us to label this negative deflection as the Nc. The morphology of this deflection was similar to that of the Nc except for the latency, although in this particular study, the peak amplitude of this early deflection was larger than the Nc. It may be construed that the negative deflection might be an artefact due to eye movements. However, we argue that it is unlikely for the following three reasons. First, the size of the stimuli presentation was adjusted so all the stimuli would fit into the central vision of infants, and they would not have to move their eyes to register the entire picture of the stimuli. Second, an object location was pseudo-randomised at each trial, and the effect of eye movement on EEG recording should be cancelled out upon averaging of the data. Third, if this were to be related to the eye movement, the incongruent condition should induce a greater peak amplitude, as the incongruent condition should require a larger eye movement than the congruent condition due to the mismatch between the eye gaze direction and the object location. For these, this deflection is unlikely to be an artefact and might reflect certain cognitive processes.

Without further investigation, it is difficult to reliably interpret this negativity. Whilst the current study does not appear to be the only case to observe a deflection in EEG before 400ms after the stimuli onset (e.g., Hoehl et al., 2008), such an early latency component is often followed by another negative deflection with a greater amplitude, which is typically interpreted as the Nc. Since the Nc has been a focus of analysis in many studies, a smaller deflection occurring before the Nc is generally not analysed, and to our knowledge, does not yet seem to have a common label. In terms of the latency, the Pb component might be closest to this early latency component. The Pb is a positive amplitude component appearing around 200-400ms after stimulus onset in central and anterior sites (Kopp & Lindenberger, 2011; Striano et al., 2006). This is thought to be linked to contextual processing and stimulus expectancy (Karrer, Karrer, Bloom, Chaney, & Davis, 1998; Karrer & Monti, 1995; Webb et al., 2005), as well as the ease of stimulus processing (Hunter & Karrer, 1993). Our result is consistent with the literature, as we observed an enhanced amplitude of this early occurring deflection for the congruent condition than the incongruent condition, and the congruent informant who consistently showed congruent gaze cues would have been easier to expect and process as compared to the incongruent informant (Köster et al., 2020). It is thus possible that this negative peak might have a similar functional property as the Pb. Nevertheless, the Pb essentially refers to a positive deflection, whereas what we observed in the current data was a negativity. Hence, it remains an open question what this early occurring negative deflection might represent. Replication of the current study as well as further examination would be needed to identify what this deflection might indicate, as this component might potentially be a useful index of neural activity which enables infants to discriminate between informants with different levels of reliability.

New questions arising from the current findings include whether informant reliability would facilitate or hinder the processing of the object about which informants give information. To test this question, for instance, the paradigm could be modified by adding another phase into the current paradigm, where only the cued object is presented after the current stimuli set. The examination of neural activity occurring in response to such object presentations would enable us to explore potential differences in infants' object processing according to the informant's reliability. The current paradigm can also be used to investigate how infants' cognitive processing of other people might modulate infants' subsequent social learning behaviour. It has been shown that socio-environmental contexts modulate a range of infants' behaviour such as imitation (Poulin-Dubois, Brooker, & Polonia, 2011; Zmyj et al., 2010), social referencing (Bazhydai et al., 2020; Stenberg, 2017), and Piagetian perseverative search task (Dunn & Bremner, 2019; Topál, Gergely, Miklósi, Erdohegyi, & Csibra, 2008). It would be interesting to explore how the neural correlates identified in the current study can explain differences in infants' subsequent behavioural task performance with previously identified reliable and unreliable informants. Such research would allow us to better understand how infants' social cognitive capacities interact with behavioural performance in general. Furthermore, this line of research investigating infants' discrimination of 'reliable' and 'unreliable' others using neural measures could place us in a position to integrate different reports in the context of selective learning behaviour during infancy, and potentially examine to what extent such infant's differential behaviour is driven by perceptions of associative, statistical regularities (Heyes, 2017a, 2017b) or reflects rudimentary understanding of other people's personality traits (Schütte et al., 2020), as the use of neural measures makes it feasible to explore infants' perceptual, attentional and cognitive processes even when overt behavioural indicators are absent.

We believe the current study has provided the first neural evidence that the reliability of informants, reflected in their gaze cue congruency, modulates infants' attentional process as well as the degree to which information is encoded. We hope that the paradigm used in the current study will be beneficial and informative for future investigation on the interplay between infant social cognition and their behaviour in socially interactive contexts.

Acknowledgement

I would like to thank all the infants and parents who took part in this study. I acknowledge that this study was completed with great help from others. The study was a secondary data analysis. The original study was designed by Dr Kirsty Dunn and Professor Gavin Bremner, and the data were acquired by Dr Szilvia Linnert with the assistance of Martina Arioli. The data had not been analysed prior to the current study. The EEG setup of this study including the MATLAB coded was provided by Barrie Usherwood.

Author contribution

Sayaka Kidby: recruitment, design of the current study (secondary analysis), data analyses, writing, review. Vincent Reid: design of the current study (secondary analysis), review. Szilvia Linnert: recruitment, data collection. Gavin Bremner: design of the original study. Kirsty Dunn: design of the original study, design of the current study (secondary analysis), review.

Chapter 4

Investigating the neural mechanism of violations of expectations in infants

Chapter Introduction

The social environment surrounding infants is dynamic. Yet, research investigating infant social cognition using neural measures typically presents infants with a series of static images to address questions of the working and development of infant social cognition. To better understand the neural processes of infant social cognition in naturalistic social situations, improving the ecological validity of experimental settings is critical. One way to do so is to incorporate dynamic stimuli, instead of static images.

In the following chapter, we present a study where we took one step closer to dynamic contexts and investigated neural substrates of 8-month-old infants' violations of expectation using multiple videos depicting different object motions. The challenges of using dynamic stimuli are discussed, which point to specified obstacles that we need to overcome before we can transition to a more dynamic and naturalistic social environment to study infant social cognition using the methods of cognitive neuroscience.

Abstract

Violations of expectation (VoE) paradigms have been used to investigate preverbal infants' understanding in our physical and psychological world (e.g. Baillargeon, 1987; Baillargeon & DeVos, 1991; Wynn, 1992), but it can also provide us with insights into how infants learn about the world and develop cognitive skills to successfully and adaptively interact with others and the surrounding environment. Despite its theoretical importance, much less is understood about the role and function of infants' ability to detect unexpected events in infant learning and cognitive development. As empirical evidence to date is varied to a large extent in terms of stimuli and neural measures used, investigation of more generic neural substrates of VoE processing can clarify the underlying cognitive processes of VoE, and might shed light on how the detection of VoE might lead to enhanced learning (Begus & Bonawitz, 2020; Stahl & Feigenson, 2015, 2019). The current study explored the neural correlates of 8-month-old infants' (*N*=15) detection and processing of unexpected events utilising event-related potentials (ERPs). Unlike prior studies, multiple dynamic video stimuli were used to maintain the ecological validity of potential findings as well as sustain infants' attention during the presentation. However, the EEG waveforms obtained contained a number of constant fluctuations, which made it challenging to identify interpretable neural activity that could indicate infants' cognitive processing of VoE. Compared with the existing evidence, the current study incites novel questions regarding how infants' comprehension of the world through VoE develops.

Introduction

From a very early stage of postnatal development, human infants show rudimentary knowledge related to the physical and social environment (Spelke & Kinzler, 2007). For instance, it has been shown that infants understand object identities and properties from as early as 2.5 months of age (Baillargeon, Spelke, & Wasserman, 1985; Dunn & Bremner, 2017; Spelke, Breinlinger, Macomber, & Jacobson, 1992), basic arithmetic rules from 5 months (Wynn, 1992), and other people's action goals from 9 months (Gergely et al., 2002; Kaduk et al., 2016; Reid et al., 2009). Such infants' knowledge has been commonly investigated using a violation of expectation (VoE) paradigm. The VoE paradigm typically presents infants with a pair of event sequences; the outcome of one is expected and the other unexpected. Infants' observation of an unexpected event is often shown to lead to longer looking (e.g. Baillargeon, 1987; Baillargeon & DeVos, 1991; Wynn, 1992) and increased pupil dilation compared to expected events (Gredebäck & Melinder, 2011). These orienting responses towards unexpected events have generally been interpreted as an indicator of infants' understanding of rules related to the physical and psychological world.

Recently, a growing body of research has investigated what VoE means in the context of infant learning, beyond its use as a research paradigm. For instance, it has been suggested that experiencing VoE plays a critical role in learning (Köster et al., 2020). This view comes from the predictive processing perspective, which originally stems from motor learning principles whereby the key purpose of the brain is to optimise the outcome of motor responses (Friston, 2005, 2010; Köster et al., 2020). In this view, the brain functions in a hierarchical structure consisting of higher and lower levels. Predictions are made at each level from lower areas, which process sensory inputs and organise motor responses, to higher areas, which are responsible for higher reasoning (FeldmanHall & Shenhav, 2019). Different levels constantly communicate with one another so that the cognitive model is updated and improves as a prediction error is detected at a lower level (Friston, 2005, 2010). VoE is considered to be a form of prediction error in this framework, and therefore serves as a signal for the brain that the present cognitive model needs refining to minimise this predictive error in the future (Clark, 2013; Köster et al., 2020).

Empirical evidence also suggests that VoE provides infants with opportunities to learn. Stahl and Feigenson (2015) presented 11-month-old infants with video clips showing either expected or unexpected events involving an object, such as a train stopping in front of a wall, or a train passing through a wall. Following the observation of events, infants were presented with two objects; one of which was in the video clips, and the other depicting a novel object. Infants who had seen an expected event showed novelty preference and chose to explore the novel object over the familiar object which they saw in the video. On the other hand, infants who had observed an unexpected event spent more time exploring the familiar object than the novel object. The orienting response towards the familiar object was interpreted as infants' motivation to learn more about the unexpected situation (Stahl & Feigenson, 2015). Moreover, the authors found that the type of object manipulation in which infants engaged matched the category of knowledge that was violated in the video clips. That is, infants who saw a 'solidity' violation (i.e., object passing through a wall) banged the object with their hand or against the highchair, whereas infants who saw a 'support' violation (i.e., object floating in the air without support) repeatedly lifted and dropped the object. The difference in infants' object manipulation was not observed when infants watched an expected event, and those who viewed an expected event even preferred to explore the novel object. These suggest that infants not only differentiated expected and unexpected events, but also understand what should happen according to their prior knowledge. Specific subsequent exploratory behaviour following VoE might indicate that infants were trying to revise the predictive model so that a predictive error could be minimised in a future similar situation.

VoE is also shown to facilitate infants' information encoding. In another experiment conducted by Stahl and Feigenson (2015), infants were presented with either an expected or unexpected event, after which they were shown that the object presented in the event ("target object") could make a squeaky sound. At the subsequent test phase, infants were presented with the target object and a novel object while the squeaky sound was played. The looking time analysis found that only infants who saw the unexpected event spent a longer time looking at the target object than the novel object when the squeaky sound was played, whereas infants who saw the expected event preferentially looked at the novel object. The authors argued that infants who experienced VoE successfully registered the object-sound association, whilst those who did not failed to do so. This led to their conclusion that the observation of VoE events enhanced infants' encoding of object property (Stahl & Feigenson, 2015). Whilst their interpretation is intriguing, the sound property of the object is not related to the knowledge violated in the event (i.e., solidity or support). Therefore, the object-sound mapping tested in this experiment provides little insights into how VoE facilitated infants' learning about the unexpected events. Furthermore, it is unclear whether infants' longer looking time towards the target object after the observation of VoE really reflects infants' learning about the object-sound association. It has been shown that the observation of VoE leads to longer looking of the target object, and when the target object is involved in an expected event, infants show a novelty preference over a target object which is more familiar (Stahl & Feigenson, 2015). Possibly, the longer looking towards the target object in the sound-object test phase might simply reflect infants' preference towards the target object after the VoE event, rather than the learning of any sound-object association. Yet, this finding does demonstrate how VoE could modulate infants' visual attentional behaviour.

Despite a growing recognition that VoE may well play a key role in infant learning, the cognitive processes, beyond attentional processes, which underlie VoE have yet to be fully understood. Identifying the neural correlates of VoE processing would shed light on the specific role of VoE in infant learning, such as how VoE might lead to enhanced explanation-seeking behaviour and better information encoding, as reported in previous behavioural research (Stahl & Feigenson, 2015). Event-related potentials (ERP) are particularly relevant in this context, as it enables us to discuss specific components in relation to particular cognitive activities (Luck, 2005). Indeed, there have been EEG and ERP studies that have utilised a VoE paradigm to investigate infants' cognitive activity (e.g., Berger, Tzur, & Posner, 2006; Kaduk et al., 2016; Kaufman, Csibra, & Johnson, 2003, 2005; Köster, Langeloh, & Hoehl, 2019; Köster, Langeloh, Michel, & Hoehl, 2021; Reid et al., 2009). Yet, these existing studies typically focused on a specific VoE event, and propose neural correlates of the processing of the specific event. The following paragraphs outline such research, summarising the neural substrates underlying the processing of specific VoE events, including N400, Nc, gamma-, and theta-band oscillatory power. The review of the extant literature leads to our conclusion that infants' generic brain responses to VoE events are not yet fully understood.

Infants' arithmetic understanding has been investigated by Berger and colleagues (2006) using a VoE paradigm with EEG. They reported infants aged 6 to 9 months showed longer looking towards unexpected events (i.e., the incorrect solution for a simple arithmetic equation) than expected events (i.e., the correct solution), replicating a previous behavioural study using the same VoE event (Wynn, 1992). They also analysed event-related potentials (ERPs) in response to correct and incorrect solutions, and compared infants' response with adults' neural response to the same stimuli. They found a large negative deflection for the incorrect solution but not for the correct solution, both in infants and adults. The latency of this deflection was later in infants (300-550ms) than adults (150-250ms), reflecting the different developmental stages of the brain (de Haan, Johnson, & Halit, 2003; DeBoer, Scott, & Nelson, 2004). These findings of this study posit the possibility of infants' rudimentary ability to process unexpected arithmetic consequences in a similar manner to adults. However, the morphology of the infants' ERP contained a number of fluctuations, making it difficult to interpret in relation to ERPs which have been linked to specific cognitive processes. Therefore, it is only possible to argue that this study demonstrated infants' ability to differentiate unexpected arithmetic sequences from expected ones. Replication of this finding would further strengthen their argument and advance our understanding of infants' cognition underlying arithmetic error detection.

Investigating infants' semantic understanding of other people's actions, Reid and colleagues (2009) utilised a VoE paradigm focusing on the N400 ERP component. They presented 7- and 9-month-old infants as well as adults with a series of static images depicting two types of eating action sequences. One of the sequences involved an expected consequence (e.g., bread reached the mouth of an adult), whereas the other ended with an unexpected consequence (e.g., bread reached an ear or forehead of an adult). They found that adults and 9-month-old infants both showed the N400 when they observed unexpected actions, indicating infants differentiated the expected and unexpected action consequences in a manner similar to that of adults (Reid et al., 2009). The N400 is a waveform which peaks to the negative polarity at approximately 400ms after the onset of a stimulus in adults, and at a later latency in infants, approximately between 600ms and 800ms after the stimulus onset (Parise & Csibra, 2012). This component has been associated with the detection of semantic (i.e., meanings) mismatch in action sequences, such as combing hair with a toothbrush (Amoruso et al., 2013). Thus, their finding suggests that infants' processing of VoE can be indexed by N400, reflecting the violation of infants' semantic knowledge (Kaduk et al., 2016; Langeloh, Buttelmann, Pauen, & Hoehl, 2020; Michel, Kaduk, & Ní Choisdealbha, 2017). Parise and Csibra (2012) also found N400 when infants detected an incongruency between a presented object and its label uttered by their parent. Consistent to the findings from Reid et al. (2009), this also suggests that N400 is sensitive to infants' detection of semantic violation. Yet, as opposed to the topics of action goals or language, to our knowledge, no investigation focusing on N400 has been conducted thus far in the context of other events used in VoE paradigms, such as object permanence, solidity or gravity (Baillargeon et al., 1985; Spelke & Kinzler, 2007; Stahl & Feigenson, 2015). Thus, it is necessary to examine whether this component can index VoE in other domains than action or lexical semantic knowledge.

Along with N400, the amplitude of the Negative Central ERP component (Nc) has been found to differ when infants observe expected and unexpected events. The Nc is a negative deflection which peaks between 400ms and 800ms following the stimuli onset, prominent over fronto-central regions (Snyder et al., 2002). It is generally thought to reflect attention allocation (Parise et al., 2008; Reynolds & Richards, 2005). Yet, the direction of amplitude difference has seen no consensus thus far. Some studies reported an enhanced amplitude of Nc for unexpected events (Kayhan, Meyer, O'Reilly, Hunnius, & Bekkering, 2019; Langeloh et al., 2020; Reynolds & Richards, 2005), while other studies reported a reduced Nc (Kaduk et al., 2016; Reid et al., 2009). Therefore, the Nc might simply reflect infants' differentiation of expected and unexpected events, and might not be the most optimal measure to detect small changes in other cognitive processes.

Another ERP component relevant to this topic is Positive Slow Wave (PSW), which has been investigated in the same paradigm where the Nc component was found (Reid et al., 2004). The PSW is typically evident after 1000ms following the stimulus onset, maximal over temporal scalp regions (Snyder et al., 2002), but early-latency PSW can become prominent around 600ms after stimulus onset and last for several-hundred milliseconds (Reid et al., 2004). This component has been thought to reflect the degree of information encoding and updating (de Haan & Nelson, 1999; Nelson, 1997; Nelson & Collins, 1991; Snyder et al., 2002). This component has not been investigated in the context of VoE, but the cognitive function that PSW is thought to represent fits well within the predictive process model. In the predictive processing model, learning is essentially an update of prior cognitive models (Friston, 2005, 2010; Köster et al., 2020). Therefore, if VoE triggers the update of the cognitive model, such model-updating processes could be indexed by PSW.

Not only ERPs but also EEG oscillatory activity has been linked to infants' observation and processing of unexpected events. Using a VoE paradigm, a neural representation of object permanence in 6-month-old infants was investigated by Kaufman and colleagues (2003). Object permanence refers to infants' understanding that an object should continue to exist when out of sight, and this has often been explored by using VoE paradigms (Baillargeon et al., 1985). Kaufman and colleagues (2003) were particularly interested in identifying the neural correlates of object representation during the time when the object is not in view ('object occlusion'). They presented 6-month-old infants with four video stimuli showing expected and unexpected appearance events, as well as expected and unexpected disappearance events. They reported infants' longer looking time towards the unexpected disappearance than the expected disappearance, but did not find any difference in looking time between the expected and unexpected appearance. They also found a difference in gamma-band oscillatory activity (30-50Hz) over right temporal channel regions between the expected and unexpected disappearance conditions, but not between the expected and unexpected appearance conditions. The authors hence argued that this gamma-band activity is likely to specifically represent the neural substrates of object occlusion (Kaufman et al., 2003). As a follow-up study, the same group of authors found an increase in gamma-band power when infants viewed gradual deletion event (i.e., another type of occlusion events), but not when they watched disintegration (i.e., disappearance) (Kaufman et al., 2005). As the stimuli in this study did not involve VoE elements, it further supports the ideal that gamma-band activity might be specifically linked to object occlusion, rather than VoE. Consistently, Southgate and colleagues (2008) also showed that gamma-band oscillation is specific to object occlusion, but not human face occlusion (Southgate et al., 2008). Nevertheless, occlusion events have frequently been used involving different objects (e.g., occluder, container, tubes) across different VoE paradigms that target understanding of different types of physical world principles (Hespos & Baillargeon, 2001a, 2001b, 2008; Luo & Baillargeon, 2005; Wang, Baillargeon, & Paterson, 2005). Therefore, an increased gamma-band oscillatory activity might perhaps index infants' VoE more in general, not strictly limited to occlusion events.

Another oscillatory rhythm that has recently been linked to infants' processing of VoE events is theta-band frequency (Köster et al., 2019, 2021). Köster and colleagues (2021) presented 9-month-old infants with a series of static images depicting four different types of VoE events. An increased theta power was found when infants viewed unexpected events as compared to expected events. Based on this, it was argued that theta-band activity is associated with infants' processing of novel information. The similar results have been reported in a study that utilised a neural entrainment technique (Köster et al., 2019). It has been known that neural oscillatory activity can be entrained to external stimuli in such a way that the frequency becomes aligned to the frequency of the stimuli presentation (Müller et al., 1998). Köster and colleagues (2019) entrained infants' brain activity to theta-band frequency (4Hz). They found that infants' theta-band oscillatory activity increased in response to the observation of unexpected events (Köster et al., 2019). Importantly, this increase in theta-band power in response to VoE was not observed in a condition where infants' brain activity was entrained to alpha-band frequency (6Hz). This comparison allows us to consider that the increase in theta-band oscillatory power following VoE is endogenously induced and reflects infants' cognitive activity, rather than driven by external factors and represents perceptual processing. Yet, whilst neural entrainment techniques are useful to dissociate perceptual and cognitive processing, the stimuli presentation is not naturalistic and lacks in ecological validity. Therefore, it is important to replicate these results using a more ecologically valid paradigm. Furthermore, the analysis by Köster and colleagues (2021) found the effect in the range of 4.5 to 5Hz, whilst a range of 3 to 5Hz is more commonly used as a theta band in infancy research (Begus & Bonawitz, 2020). Therefore, it is also important to examine whether this specific "upper" bound of theta frequency is more sensitive to the processing of VoE than a conventional theta frequency band.

To sum, many of the existing studies have investigated specific neural correlates of a specific type of VoE, such as occlusion and action goals. Nonetheless, it is less clear whether these neural substrates reported so far reflect the processing of specific type of VoE, or a more generic neural response to unexpected events. A few recent studies by Köster and colleagues (2009, 2021) have used multiple VoE event presentations and found that the upper end of theta-band oscillation might index infants' processing of VoE generally. Despite this, replication of the effect is important, particularly with more naturalistic paradigms than their paradigm involving the use of static or flickering images.

The current study aims to explore the neural substrates of infants' general VoE processing across various event types, according to the proposal that VoE generally triggers infants' cognitive model improvement and facilitates learning (Köster et al., 2020). Based on prior studies, it is hypothesised that the following three ERP components could be observed as a response to unexpected motion sequences; an enhanced Negative Central component (Nc), an N400, a larger peak amplitude of Positive Slow Wave (PSW), as well as increased theta- and gamma- band oscillatory activity (e.g., Kaufman et al., 2003, 2005; Köster et al., 2019, 2021; Michel et al, 2017; Parise & Csibra, 2012; Reid et al., 2009). These ERP components would indicate whether infants' VoE involves their enhanced attentional allocation to unexpected events, detection of semantic incongruency, or information updating. By examining theta and gamma oscillatory activity, we aim to examine if they could be a neural marker of general VoE in infants, not specific to certain types of VoE events (Kaufman et al., 2003, 2005; Köster et al, 2019, 2021; Southgate, 2008).

Methods

Participants

The final sample consisted of 15 infants (6 females), with the average age being 7 months and 22 days (ranging from 227 days to 251 days, SD_{age} = 6.70 days). All infants were born full term (37-41 weeks) and in the normal range for birthweight (over 2500 grams). The age group was selected based on the existing evidence that 8-month-old infants should understand the underlying object principals depicted in the stimuli used in the current study, while younger infants may fail to recognise impossible events (e.g., Hespos & Baillargeon, 2001). An additional 36 infants were tested but were excluded from the final sample due to fussiness (*n* = 2), not having completed the minimum number of trials required to conduct the data averaging of the standard ERP analysis (*n* = 32), or a technical or experimental error (*n* = 2). All infants were recruited through phone calls and emails from a database consisting of parents who expressed an interest in taking part in psychological research.

Prior to recruitment, ethical approval regarding the study procedure (e.g., recruitment, research methods and data handling throughout the study) was sought and obtained from the Lancaster University Faculty of Science and Technology Research Ethics Committee. All the parents whose child participated in the current study were given a full description about the study from the experimenter, and gave a written informed consent on behalf of their child.

Stimuli

The video stimuli were filmed at Lancaster University. The stimuli consisted of seven pairs of object motions, comprising a motion consistent with infants' expectation (the expected condition) and the other motion violating their expectation (the unexpected condition). The presentation of multiple different object motions, rather than focusing on one single object motion, was critical to address the main question of the current study, and was also considered to be optimal to avoid potential attenuation of the VoE effect in infant ERPs due to the repetition of the same stimulus presented, which could mitigate the ERP effects (Stets, Burt, & Reid, 2013).

Videos were edited to remove sound and to restrict the view only to the target motion (i.e., no faces were visible in the video clip). This was to prevent infants' distraction during the critical moment of object motions (i.e., when the video clip shows the expected or unexpected motion), and avoid any changes in stimulus processing that might result from facial or vocal ostensive cues from the actor in the video. Each pair of the videos differed in length due to naturalistic variations in the actor's manual timing in their movements. The speed and the length between object motion were adjusted after the recording to ensure we obtained a sufficient number of EEG trials with limited data recording time, while maintaining the ecological validity of the stimuli as much as possible. Therefore, there were minor variations in terms of the length of each video clips (M_{duration} = 5.21 seconds, SD_{duration} = 1.46 seconds). However, since the data were analysed according to the object motion being expected or unexpected, the effect of such variations across video clips was considered negligible if any.

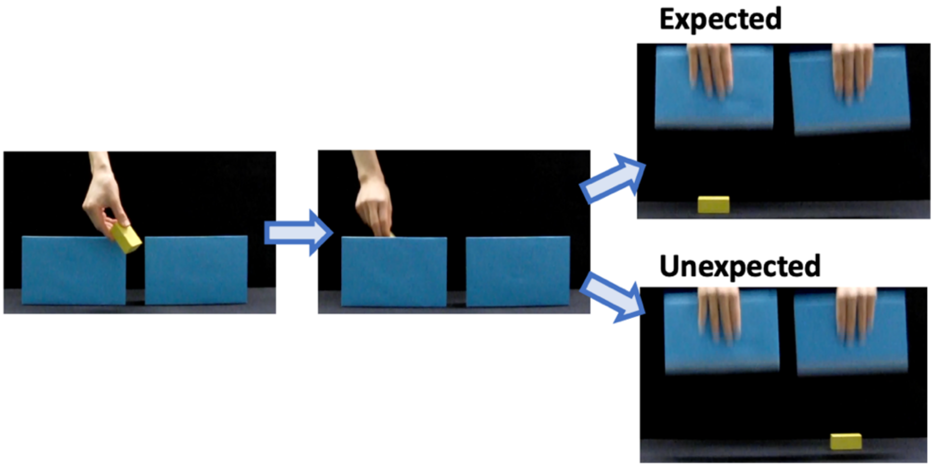
The presentation order of the videos was determined using pseudo-randomisation, in terms of condition (expected or unexpected), as well as video categories (i.e., which video was presented). More specifically, no more than three video clips showing the same condition (expected or unexpected) or the same type of motions (e.g., "box behind the screen", "unexpected appearance") were consecutively presented.

Each image subtended a visual angle of approximately 13° horizontally and 15.5° degrees vertically on a 22-inch screen (resolution 1350 x 1080 pixels). The distance between the monitor and the infant was approximately 70cm.

Following is a list of seven pairs of object motions, including the total of 14 different videos used in the current study.

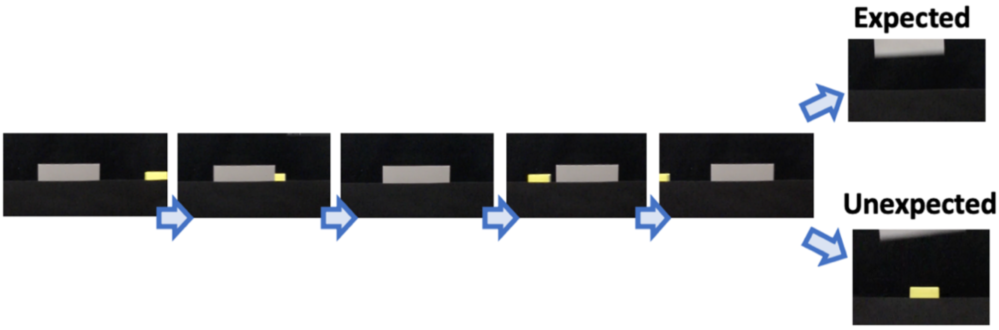
(a) **Temporal continuity (box behind the screen) event**. The stimuli consisted of two identical blue screens made with a cardboard box (14 x 21cm) that occluded a yellow wooden rectangular box (2.7 x 6 x 2.7cm). Infants saw an event where the actor placed the two identical screens on the stage, approximately 3 cm apart so that infants could see they were not connected. The actor showed several different sides of the yellow box, and hid it behind the left screen. The actor then lifted both screens simultaneously to reveal the object. In the expected condition, the object was still behind the left screen, while in the unexpected condition, the object was behind the right screen. Each condition took 7 seconds.

Figure 4-1. Illustrations of the stimulus (box behind the screen).



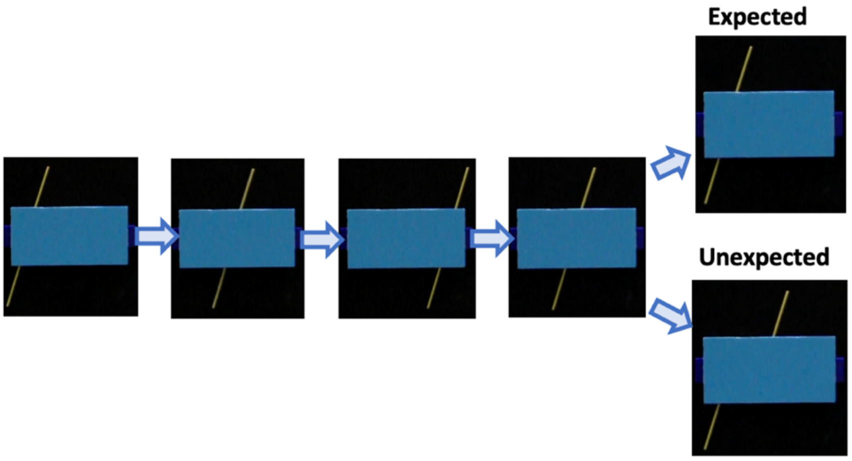
(b) **Temporal continuity (unexpected appearance) event**. The stimuli involved a yellow box (5.5 x 3 x 2.6cm) and a grey screen (10.5 x 15.1cm) placed in front of a black board as a background. Infants saw a yellow box moving from the right-hand side. The box became occluded and hidden behind the grey screen, and then re-appeared from behind the screen toward the left-hand side. In the expected condition, infants saw nothing behind the screen, whilst in the unexpected condition, infants saw a box resting behind the screen when it should not be there. Each sequence lasted 6 seconds.

Figure 4-2. Illustrations of the stimulus (unexpected appearance).



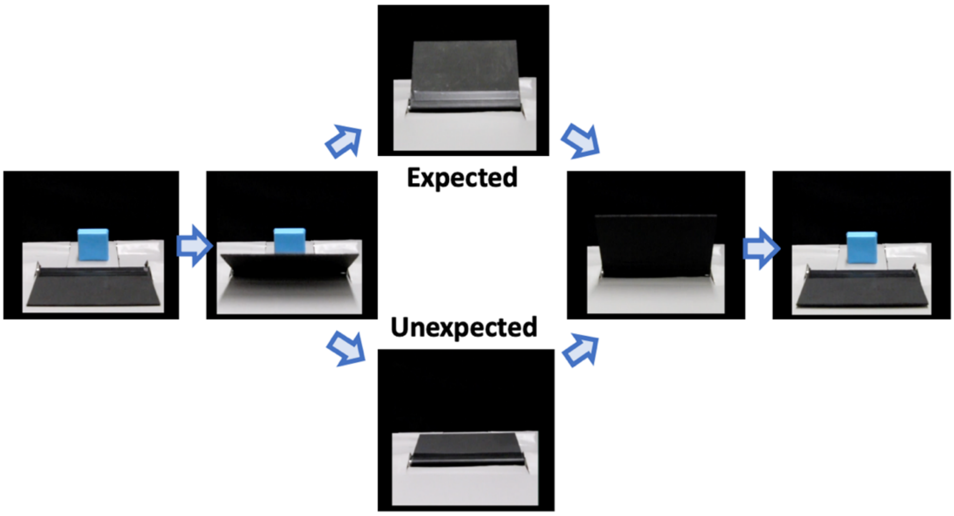
(c) Spatial continuity (broken rod) event. The stimuli involved a yellow rod (which is actually two short rods, with each visible portion being approximately 7cm long) and a blue screen (16 x 8cm) which was hiding the middle part of the rod. In the expected condition, infants saw the rod(s) moving behind the screen from the right-hand side to the left-hand side, taking around 1 second, and then back to the right-hand side at the same speed, with only the top and bottom edges being visible (i.e., the middle part was hidden behind the screen). In the expected condition, infants saw this trajectory (from right to left and back to the right) twice. In the unexpected condition, infants saw this normal trajectory once, and saw only the top visible potion of the rod moving, while the bottom remain stationary. Each condition took 4 seconds.

Figure 4-3. Illustrations of the stimulus (broken rod).



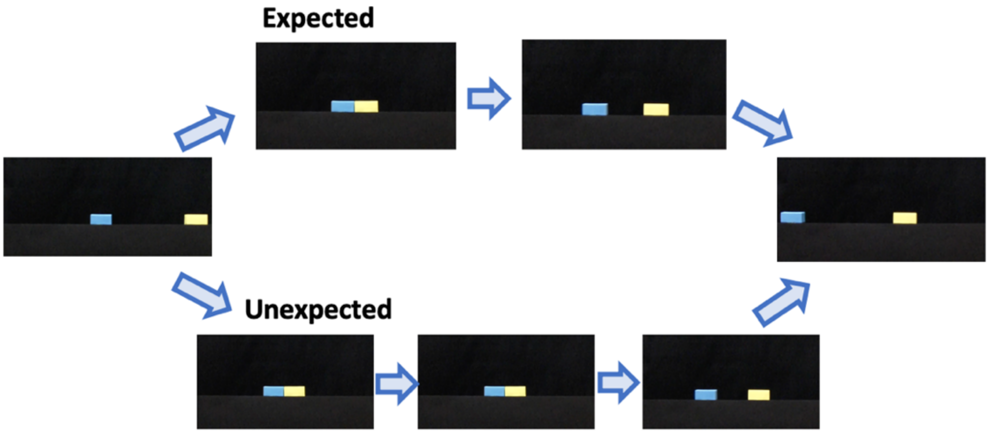
(d) Object permanence (drawbridge) event. The stimuli consisted of a black screen (15 x 9.5cm) that rotated back and forth either 120° or 180°, and a blue block (4.3 x 4.3 x 2cm) as an occluder of the screen. After infants saw a blue block sitting in the path of the rotating screen, in the expected condition, infants saw a screen rotated for 2 seconds away from the infant for 120°, occluded by the block, and then paused flat for 1second. In the unexpected condition, infants saw a screen rotated away from the infant and stopped flat (i.e., rotated 180° away from the infant) despite the solid block placed in the pathway. Each sequence lasted 5 seconds.

Figure 4-4. Illustrations of the stimulus (drawbridge).



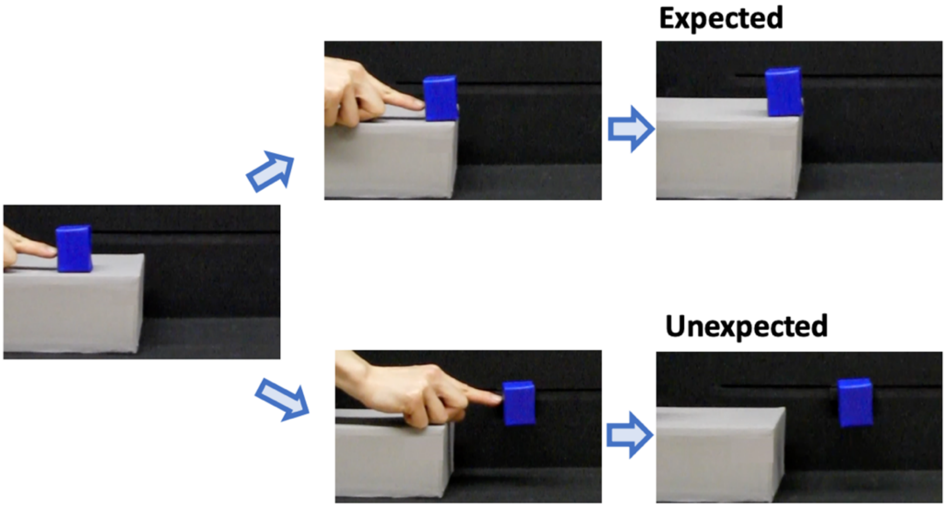
(e) Causality (contact) event. The stimuli involved two rectangle blocks (one yellow, one blue, each 3.5 x 2.6 x 2.6cm), each of which moved for 18.1 cm in such a way that one block collides with the other. In the expected condition, infants saw one box hitting the other box, which then would move immediately after being hit, whereas in the unexpected condition, the left object only started to move with a delay of 1 second after being hit. Each sequence took 3 seconds.

Figure 4-5. Illustrations of the stimulus (causality).



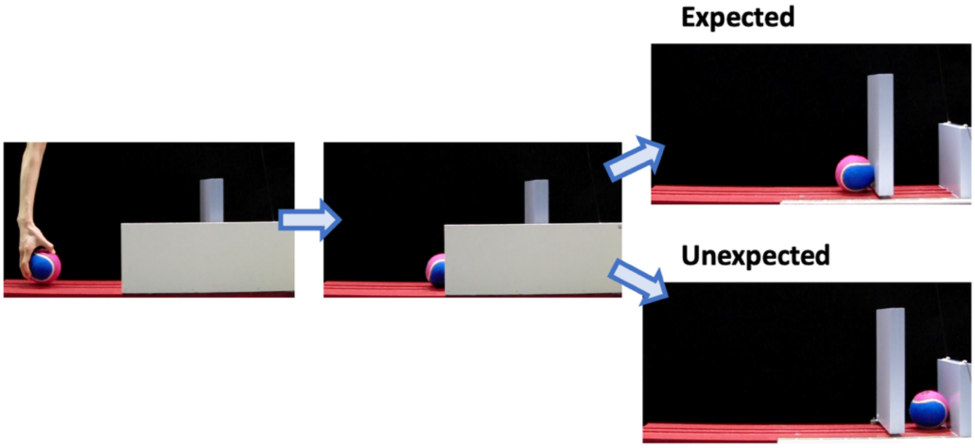
(f) Gravity event. Infants saw a black stage and a small grey box (16 x 5.5 x 5.5cm) with a smaller blue box (3.5 x 2 x 2cm) on it. In the expected condition, infants saw a blue box slowly pushed toward the edge of the grey box, which stayed at the edge, balanced on the stage without falling off. In the unexpected condition, infants saw a blue box slowly pushed toward and beyond the edge of the grey box without falling, and floating in the air without having any contact with the grey box. Each event lasted 4.5 seconds.

Figure 4-6. Illustrations of the stimulus (gravity).



(g) Solidity (ball and barrier) event. The stimuli involved a stage, two solid boards with an occluder attached, as well as a ball (9cm diameter). Infants first saw an empty stage with only one blue board placed at the right side of the stage. The other blue board was brought down to near the centre of the stage by an actor (face not visible) showing the board was solid with no ball. Infants then saw a ball rolling down towards the blue boards. In the expected condition, infants saw a ball occluded at the board in the middle. In the unexpected condition, infants saw a ball occluded at the board placed at the far end of the stage, as if it had passed through the solid board in the middle. Each event took 7 seconds.

Figure 4-7. Illustrations of the stimulus (ball and barrier).



Procedure and apparatus

An experimental session commenced in a room specifically designed for the current study. Prior to the commencement of the testing session, parents were instructed not to allow the infant to move excessively and not to engage in communication with the infant during the stimuli presentation. Such precaution was taken to ensure that parents did not influence infants' responses to the stimuli, and also to prevent EEG data from being contaminated by artefacts.

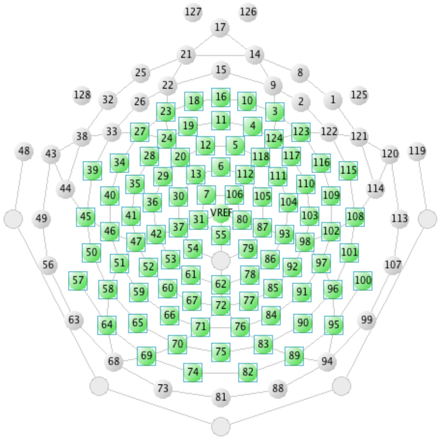
Infants sat on their parent's lap in front of a screen, and viewed both expected and unexpected sequences, contributing data to both conditions. In order to minimise possible distractions for infants, light conditions in the laboratory were kept low, and a curtain was placed to separate the experimenter and EEG equipment from the infant. A camcorder was placed at the top of the screen view which presented the video stimuli, and it captured the upper body of the infant (up to their chest). This video feed was used by the experimenter, who sat behind the curtain out of the infant's view, to decide whether to engage with the parent for assistance or respond to issues during recording, such as the infant pulling at the electrode cables on the EEG cap. Using the camcorder, infant behaviour was concurrently recorded during the EEG data recording. The video recordings were used to code infants' visual attention towards the stimuli on a trial-by-trial basis offline after the infant had left the laboratory.

The video presentation was terminated when infants paid no more attention to the video. Hence, not every infant saw the same number of each motion sequence. Overall, infants watched 76 video clips on average (ranging from 34 to 133, $SD_{videos} = 31.51$) before they failed to reliably attend to the screen. For the final analyses (i.e., grand averaging), 208 out of 570 trials (36%) for the expected condition, and 199 out of 566 trials (34%) for the unexpected condition were included. The average numbers of trials included in the final analyses per participant were 14 ($SD_{trials} = 8.53$) for the expected condition and 13 ($SD_{trials} = 7.18$) for the unexpected condition. The number of trials included in the analyses as well as the number of trials which infants watched in total did not differ across conditions ($p = .207$, $p = .389$ respectively).

EEG Recordings and analyses

Infants' electrical brain activity was measured by a 128-channel Geodesic Sensor Net while infants watched the video stimuli. EEG data were amplified with an EGI Net Amps 300 amplifier with a sampling rate of 500Hz (HCGSN 130, EGI, Eugene, OR, USA). The four channels (channels 125, 126, 127, 128) were disconnected and not used in the current study to avoid excessive discomfort for infant participants, having electrodes close to their eyes. Therefore, the data for the current study were collected by 124 channels in total. Furthermore, the electrodes from the most anterior and posterior areas were also excluded from the final analysis due to a high noise level caused by poor contact with the scalp, which is a common data management process in infant EEG studies (Bakker et al., 2016; Gredebäck et al., 2015, 2010). Therefore, 38 electrodes in total were excluded from the final analyses. The final analyses were conducted over the data collected from 90 channels (see Figure 4-8).

Figure 4-8. Channels included in the final examinations.



EEG was referenced online to the vertex electrode (Cz), with an analogue bandpass filter being applied (0.1 to 100 Hz) when storing the data for subsequent offline analyses. When the expected or unexpected outcome was revealed in a video, a signal (trigger) was sent to the EGI system via a MATLAB script to enable ERP analyses. For this, the timing at which a trigger should be sent was manually extracted frame-by-frame using iMovie (30 frames per second). The accuracy of these timings was assessed by comparing the timing of the inputs in the MATLAB with the timing of outputs in the NetStation. Out of 220 triggers sent in total, the mean difference observed between the MATLAB inputs and the NetStation outputs was 9.64 milliseconds, with the standard deviation being 2.54 milliseconds (ranging from -3 milliseconds to 13 milliseconds). The difference in timing was used to adjust the ERP latency.

Initial ERP data processing was performed using NetStation 4.5.4 Waveform Tools. The data were first offline-filtered using 0.3Hz - 60Hz bandpass, which then were segmented so each segment would include a time window from 1000ms before the outcome (expected or unexpected) was revealed to 2200ms after the outcome was revealed. This segmentation was based on the literature which reported that components and oscillatory activity of focus had been observed within 1200ms after the stimuli onset (Kaufman et al., 2005; Parise et al., 2008; Reid et al., 2004; Senju et al., 2006; Striano et al., 2006). The segmentation in the current study was longer than the time window of interest for the planned ERP and time-frequency analyses, as the wavelet transform necessary to conduct a time-frequency analysis could distort the edges of the EEG segments. After the segmentation, the automatic artefact detection tool was applied to segmented EEG data (i.e., trial), which marked a channel in each trial as “bad” when the amplitude exceeded the set threshold of +/- 200 microvolts for the entire segment, with a moving average of 80ms. This automatic artefact detection was manually and visually inspected afterwards to improve the accuracy of artefact detection, especially for detecting the artefact due to eye blinks, eye movement and body movements (for further information on infant ERP methods, see Hoehl & Wahl, 2012). The video data recorded during the EEG recording were also used for the manual rejection of EEG trials, to remove contamination due to infant body and eye movement.

Based on the manual artefact detection, trials which included more than 12 channels (i.e., more than 10% of the channels used for data acquisition) with excessive noise levels were excluded for the following steps of the analysis. At this time point, the data from 36 participants out of 51 participants who contributed less than five usable trials were excluded from the subsequent analyses. This data inclusion criteria of five usable trials per condition has been used in previous studies, which showed that a reliable and interpretable analyses on infant ERP data can be performed with 3-7 usable trials per condition (Kaduk et al., 2016, 2013; Stets & Reid, 2011).

The channels marked as noisy in the final sample of 15 participants were interpolated by using the average of surrounding electrodes. Following the interpolation, the channels were re-referenced to the average of all electrodes. Following this, the baseline collection was performed, with the baseline being the time window between -200ms and 0 relative to the timing at which the expected or unexpected outcome was displayed. Afterwards, an individual average for each participant was calculated. The grand-average waveform using the 15 participants' averages was then computed.

The electrode sites to be investigated in the following analyses were determined based on the existing literature. Namely, the Nc was examined over fronto-central electrode sites (Hoehl et al., 2008; Reid & Striano, 2005; Wahl et al., 2013), the PSW over fronto-temporal regions (Reid et al., 2004), the N400 over parietal regions (Parise & Csibra, 2012), theta oscillatory activity over the entire scalp region (Köster et al., 2021), and gamma oscillatory activity over temporal regions (Kampis, Parise, Csibra, & Kovács, 2015; Kaufman et al., 2003, 2005). The specific channels included in each analysis were summarised in Table 4-1.

Table 4-1. Channels included in analyses of ERPs and oscillatory activities.

	Channel regions	Channel numbers (128-channel Geodesic Sensor Net)
Nc	Fronto-central	4 5 6 10 11 12 13 16 18 19 20 24 112 118 124
PSW	Fronto-temporal	24 27 28 34 116 117 123 124
N400	Parietal	52 53 58 59 60 61 62 67 72 77 78 85 86 91 92 96
Theta	Entire scalp	1 – 124 (except for the excluded channels)
Gamma	Temporal	39 45 50 57 58 64 95 96 100 101 108 115

Results

The Negative Central component

The Negative Central component (Nc) was examined over fronto-central electrode regions, within a time window between 400ms and 600ms after the expected or unexpected object motion was displayed in each video clip.

To examine the data quality, the power spectrum was generated using the EEGLAB function, namely the “precompute channels measures” included in the STUDY setting with the default specifications for computing the power (i.e., 'specmode', 'fft', 'logtrials', 'off'). The examination of the power spectrum indicated that the data were contaminated at approximately 50Hz (see Figure4-9), most likely due to noise from the amplifier. Since ERP waveforms typically consist of frequency range between approximately 0.1 Hz and 30Hz (Luck, 2005), it was decided to apply an offline high-pass filter at 30Hz to reduce the noise level.

Figure 4-9. Power spectrum of the EEG data recorded over fronto-central channel regions.

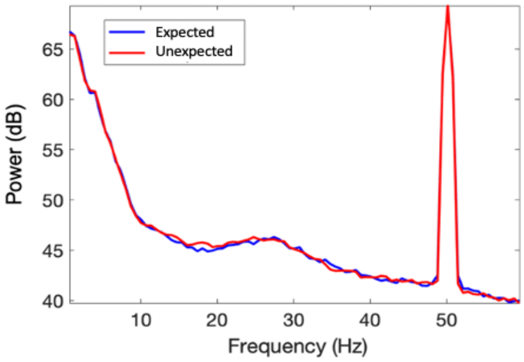
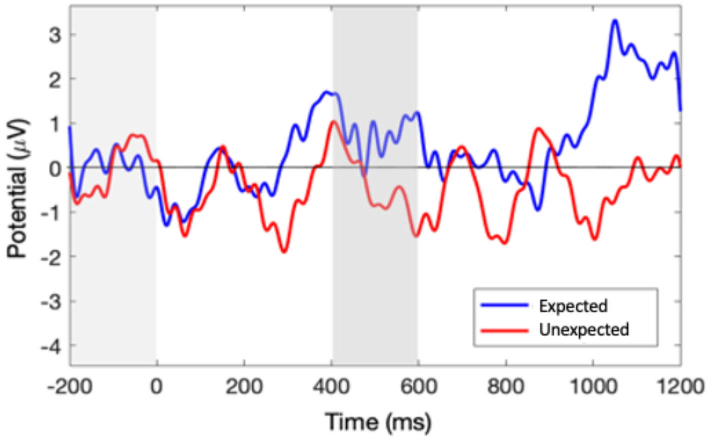


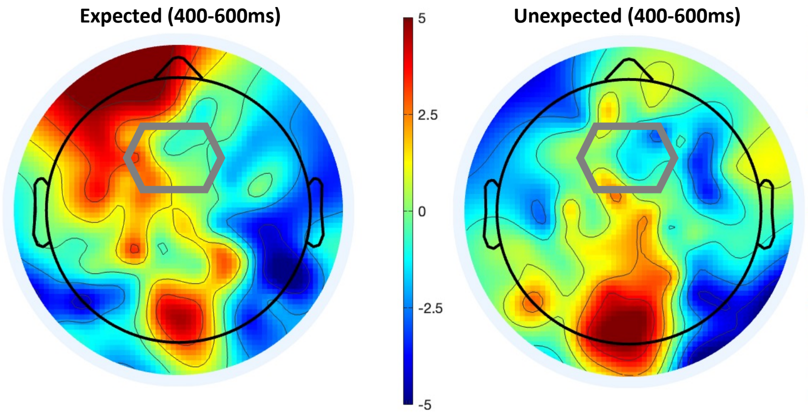
Figure 4-10 shows grand average waveforms after the high-pass filter at 30Hz was applied. Figure 4-11 shows the average scalp distribution of the grand average waveforms, within a time frame between 400ms and 600ms, for the expected and unexpected conditions. Visually inspecting the waveforms as well as the scalp distribution plot, we concluded there was no interpretable difference in the Nc component. A paired sample *t*-test also suggested no significant difference between conditions ($t(13) = 1.13$, $p = .28$, Cohen's $d = 0.29$) ($M_{expected} = 1.05$, $SD_{expected} = 4.14$ [μV], $M_{unexpected} = -.503$, $SD_{unexpected} = 4.86$ [μV]).

Figure 4-10. Grand-average waveforms (N = 15). ERPs elicited for the expected (blue) and unexpected (red) conditions.



Note. The time = 0 refers to the time at which the expected or unexpected object motion was displayed on the screen. The darker grey shading indicates the time window used to examine the Nc (400-600ms), and the lighter grey area marks the baseline period (-200ms to 0ms relative to the stimuli onset).

Figure 4-11. Average scalp distribution of the grand average ERPs (N = 15) between 400ms and 600ms for the expected (left) and unexpected (right) conditions.



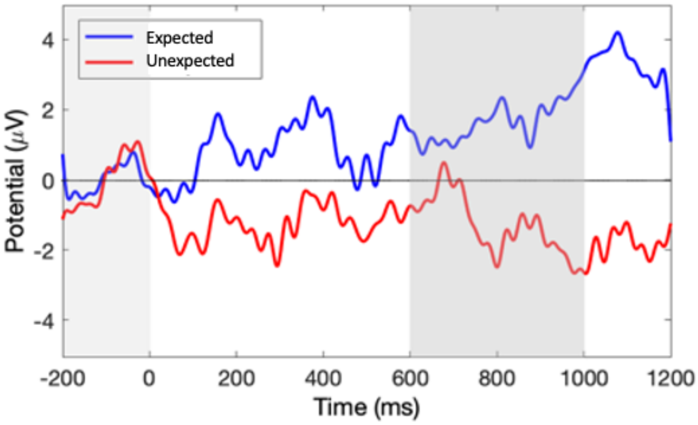
Note. The grey area represents the fronto-central electrode regions examined.

The Positive Slow Wave

The Positive Slow Wave (PSW) was inspected over fronto-temporal electrode sites, with a focus on a time window between 600ms and 1000ms after the expected or unexpected object motion was displayed on the screen in each video clip. As with the Nc, the power spectrum of the data suggested a high level of noise at approximately 50Hz. Hence, the data were high-pass filtered at 30Hz to reduce the noise level while maintaining the necessary frequency range to examine ERPs (Luck, 2005).

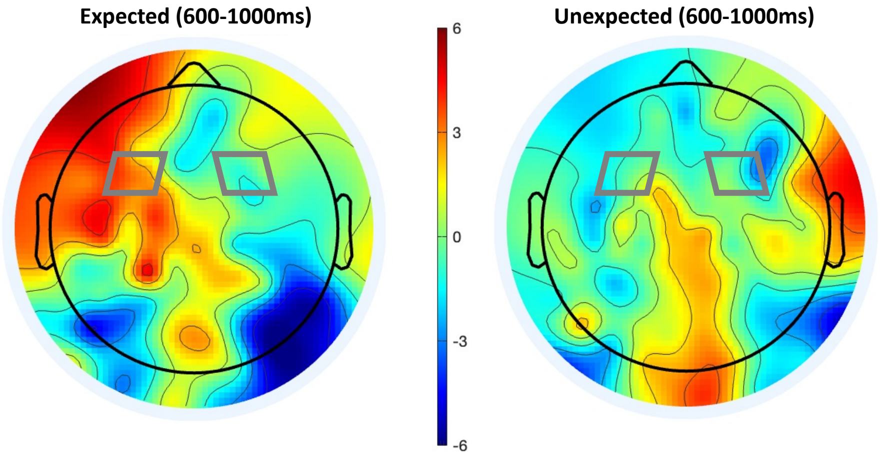
Figure 4-12 shows grand average waveforms after the high-pass filter at 30Hz was applied, and Figure 4-13 depicts the average scalp distribution of the grand average waveform. The examination focused on the time interval between 600ms and 1000ms after the expected or unexpected object motion was displayed. A paired sample *t*-test suggested that the mean amplitude of the PSW within the time frame of interest was larger for the expected condition ($M_{expected}$ = 2.40, $SD_{expected}$ = 3.79 [μV]) than the unexpected condition ($M_{unexpected}$ = -1.38, $SD_{unexpected}$ = 3.60 [μV]) ($t(13)$= 2.26, p = .04, Cohen's d = 0.58). However, the fluctuations in amplitude seen in the waveforms suggested that the quality of the data was likely not sufficient to produce a reliable and interpretable ERP with a clear morphology.

Figure 4-12. Grand-average waveforms (N = 15) ERPs elicited for the expected (blue) and unexpected (red) conditions.



Note. The time = 0 refers to the time at which the expected or unexpected object motion was displayed on the screen. The darker grey shading indicates the time window used to examine the PSW (600-1000ms), and the lighter grey area marks the baseline period (-200ms to 0ms relative to the stimuli onset).

Figure 4-13. Average scalp distribution of the grand average ERPs (N = 15) between 600ms and 1000ms for the expected (left) and unexpected (right) conditions.



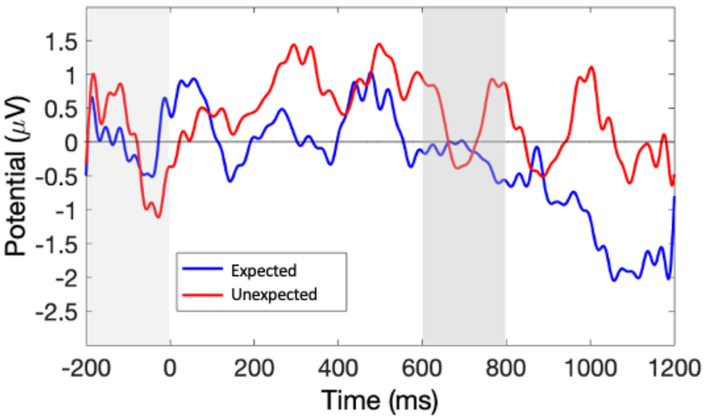
Note. The grey area represents the fronto-central electrode regions examined.

The N400

The N400 was examined over parietal electrode regions within a time window from 400ms to 600ms after the expected or unexpected object motion was revealed in each video clip. As was the case with the Nc and PSW, the power spectrum of the data indicated a high level of noise at approximately 50Hz. Hence, the data were high-pass filtered offline at 30Hz to decrease the noise level while ensuring the necessary frequency range to examine ERPs should remain (Luck, 2005).

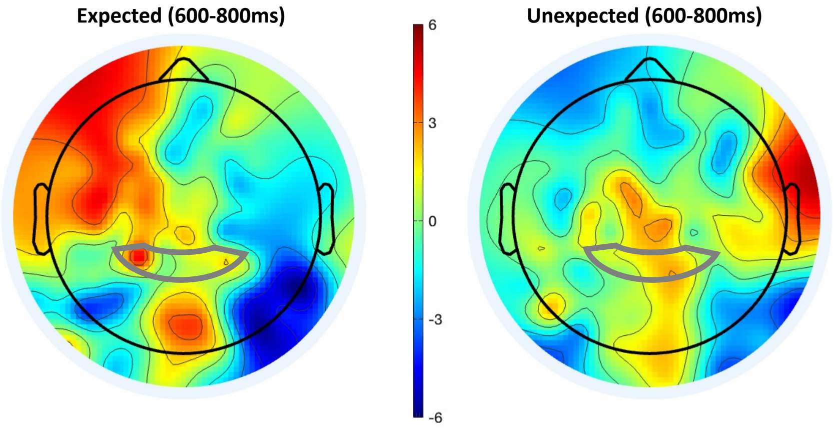
Grand average waveforms after the high-pass filter at 30Hz was applied are shown in Figure 4-14, and the average scalp distribution of the grand average waveform is shown in Figure 4-15. The data were examined over a timeframe between 600ms and 800ms after the expected or unexpected object motion was displayed on the monitor. Upon visually inspecting the waveforms as well as the scalp distribution, we concluded that there was no interpretable difference regarding the N400 component. A paired sample *t*-test also indicated there was no significant difference between conditions ($t(13)=.17$, $p = .87$, Cohen's $d = 0.04$) ($M_{expected} = -0.243$, $SD_{expected} = 2.11$ [μV], $M_{unexpected} = -.36$, $SD_{unexpected} = 2.00$ [μV]).

Figure 4-14. Grand-average waveforms (N = 15) ERPs elicited for the expected (blue) and unexpected (red) conditions.



Note. The time = 0 refers to the time at which the expected or unexpected object motion was revealed on the screen. The darker grey shading indicates the time window used to examine the N400 (600-800ms), and the lighter grey area marks the baseline period (-200ms to 0ms relative to the stimuli onset).

Figure 4-15. Average scalp distribution of the grand average ERPs (N = 15) between 600ms and 800ms for the expected (left) and unexpected (right) conditions.



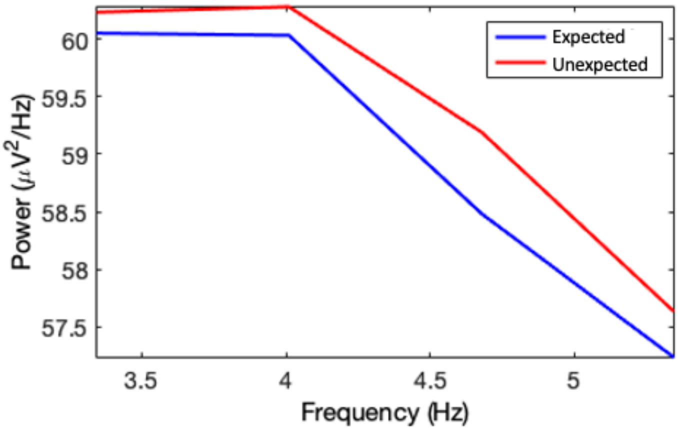
Note. The grey area represents the fronto-central electrode regions examined.

Theta- and gamma-band oscillatory activity

It was originally planned to perform a time-frequency analysis to examine bursts in theta- and gamma-band oscillatory activity as an index of infants' violation of expectations. However, the examinations of ERPs suggested that there might be a latency jitter in the data, which could cancel out the ERP effect (Cohen, 2014). Hence, we conducted a frequency analysis instead of a time-frequency analysis, without focusing on the temporal aspect of the data. Based on the data quality seen in the ERP analyses, the analysis was solely for exploratory and illustrative purposes, and hence, statistical assessment was not performed on these data.

Theta-band activity was examined over the entire scalp regions, following Köster et al. (2021). The power spectrum of the data was extracted between 3 and 6 Hz, using the EEGLAB function (Begus & Bonawitz, 2020). The visual inspection of the data, as shown in Figure 4-16, suggested that theta-oscillatory power was higher in the unexpected condition than the expected condition. Nevertheless, as noted, the quality of the data was considered insufficient to argue for the effect.

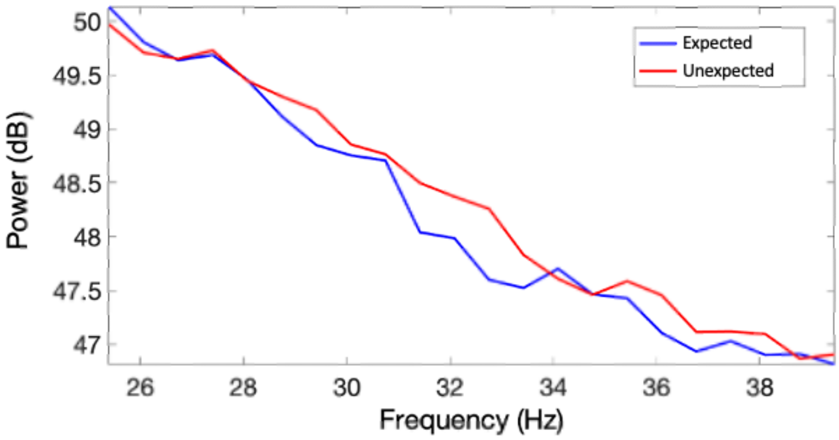
Figure 4-16. Power spectrum of the EEG data recorded over the whole brain regions for the expected (blue) and unexpected (red) conditions, showing the oscillatory activity in the theta band (3 to 6Hz).



With regards to the gamma oscillatory activity, the examination was conducted over temporal electrode sites (Kaufman et al., 2003). The power spectrum of the data in each analysis was generated using EEGLAB. A gamma frequency band has typically been defined between 20 Hz and 60Hz (Kaufman et al., 2005), but studies utilising a similar paradigm and stimuli set to those of the current study focused on the lower-frequency range in the gamma band, limiting the highest frequency lower than 60Hz. For example, prior studies have defined infant gamma range as between 25 and 35 Hz (Kampis et al., 2015), or between 32 and 48Hz (Reid et al., 2007). Hence, the frequency analysis of the current study focused on the lower and narrower gamma band (25 and 40 Hz). This enabled us to appropriately capture the neural responses in the gamma band to the stimuli, while excluding the excessive level of noise at approximately 50Hz, which was found in the previous ERP analyses.

Figure 4-17 illustrates the power spectrum between 25 and 40Hz (infant gamma band). Visually inspecting the data, we concluded that there was no significant difference in gamma band oscillatory activities across conditions.

Figure 4-17. Power spectrum of the EEG data recorded over temporal channel regions for the expected (blue) and unexpected (red) conditions, showing the oscillatory activity in the gamma band (25 to 40Hz).



Discussion

The current study aimed to identify neural correlates of infants' detection and processing of violation of expectation (VoE) in multiple event sequences. Our analyses focused on three event-related potential (ERP) components, the Nc, PSW, and N400, as well as theta- and gamma-band oscillations, all of which have been associated with infants' processing of unexpected events (Kaufman et al., 2003; Köster et al., 2021; Langeloh et al., 2020; Parise & Csibra, 2012; Reid et al., 2009; Reid et al., 2004). The aim of the investigation was to shed light on the role of VoE in infant learning (Stahl & Feigenson, 2015, 2019). Whilst there has been a body of research which investigated neural processing of unexpected event sequences and suggested several neural correlates (e.g., Kaufman et al., 2005; Köster et al., 2019, 2021; Michel, Kaduk, Choisdealbha, & Reid, 2017; Parise & Csibra, 2012; Reid et al., 2009), many of these studies typically focused on one specific domain of events, such as occlusion or action sequence (e.g., Kaduk et al., 2016; Kaufman et al., 2003; Reid et al., 2009). Therefore, what cognitive activity the processing of VoE involves in general had not been explored yet.

Overall, the current data did not show clear ERP components. Whilst there were some waveform morphologies which may indicate ERP components, we took a conservative approach and argue that the question needs to be re-addressed in future investigation with a modified design, as will be detailed in the following paragraphs. Our concern for reporting ERPs in our current data arises mainly from a number of deflections evident throughout the time windows of interest. Constant fluctuations in a waveform makes it challenging to interpret a neural activity pattern observed as an ERP component occurring in response to the stimuli, as opposed to merely a part of continuous deflections. Whilst there are prior studies that have reported ERPs from the waveforms involving as many fluctuations as our current data (e.g. Berger et al., 2006; Friedrich & Friederici, 2004, 2005a, 2006), we considered it to be appropriate not to claim any ERPs at this stage. Based on the inability to interpret the ERP components, we decided not to conduct any statistical assessment on the oscillatory activity in theta and gamma bands occurring during the stimuli presentation. The data did indicate that a theta-band oscillatory power might index infants' differential processing of expected and unexpected events, whilst future investigation is necessary to confirm this. The cognitive function that theta-band oscillation reflects in infants is yet to be fully understood, but it has been posited that it might be a marker of infants' building semantic knowledge structure (Köster et al., 2017). Should theta-band oscillatory rhythms represent infants' VoE, that would support the idea that VoE plays a key role in infant learning (Begus & Bonawitz, 2020).

The absence of clear ERP components might have stemmed from our study design involving dynamic video stimuli presenting various ranges of VoE events. Such diverse stimuli might have resulted in latency jitter across trials, which is known to diminish ERP effects (Cohen, 2014). The use of different types of object motions was desirable in order to sustain infants' attention to ensure we obtain a sufficient number of trials for reliable EEG analyses (Stets et al., 2013). However, a wide range of VoE events presented might have meant that some events were more salient and easier to process, while others were relatively complex and required more time to register. Indeed, the video clips used differed from one another in several ways. For instance, VoE events used in the current study differed in terms of the perceptual input. For example, the broken rod event simply consisted of a screen and a rod, but the ball and barrier event included hands, a ball, and two screens. More items included in the event might have resulted in a greater cognitive demand to process the trajectory of the event, perhaps contributing to a variation of neural response patterns across different VoE events. Furthermore, some events involved one type of VoE, whereas other events may have simultaneously demonstrated multiple VoE consequences. For instance, whilst the unexpected appearance/disappearance event only demonstrated unexpected appearance of an object, the box behind the screen event involved both unexpected appearance as well as unexpected disappearance. Multiple events occurring simultaneously are likely to take longer for young infants to process and register than a single event. Taken together, it is likely that cognitive and memory load to process a presented event might differ across different object motions. Along with the latency jitter, the use of dynamic video stimuli, instead of static images, might also have reduced the quality of data due to infants' eye movement during the baseline period. Typically, the use of fixation stimulus (e.g., a cross or circle) is used to control infants' saccadic patterns in the baseline period, as variations in saccadic movements could distort the data within the timeframe of interest. However, our EEG analyses were time-locked to the moment when the expected or unexpected object motion was revealed in each video. Therefore, it was not possible to present a fixation stimulus before the time-locking point. Hence, there might have been differences in saccadic movements in baseline across video types, interfering with the data quality. These issues need to be addressed in future studies by decreasing the number of event types used as stimuli to avoid latency jitter as well as variations in saccadic movements in the baseline period. Alternatively, a larger sample size might help future investigation so separate examinations for each VoE video types could be reliably performed.

Whilst further investigation is necessary, the latency jitter suggested in the current study might indicate that infants' neural processing of VoE might differ across events. This would suggest that infants might process different VoE events in a different way. This would be interesting to examine, as whilst all of the events used in the current study have been categorised as "unexpected" in the prior literature, they might be processed differently by infants with developing cognitive skills. If this was the case, what VoE means in the context of infants' learning might differ according to their cognitive developmental stages. Prior studies have found evidence that infants processing of the physical world might work in a grading scale, where some events are easier to process than others. For instance, it has been shown that infants can understand from around 3 months of age that the top of the object should be visible even when the screen is placed in front of the object to occlude its bottom part, if the height of an object is taller than the screen (Baillargeon & DeVos, 1991; Baillargeon & Graber, 1987). Yet, when the container is used instead of the screen to hide the object, infants younger than 7 months of age fail to show the evidence of such understanding (e.g., Hespos & Baillargeon, 2001a, 2001b, 2008; Luo & Baillargeon, 2005; Wang, Baillargeon, & Paterson, 2005). More specifically, in a study conducted by Hespos and Baillargeon (2001a), infants saw a short or tall piece of plastic tube being lowered into a container. In the tall container event, the container was tall enough to fully hold the whole part of the plastic tube, whereas in the short container event, the container was shorter than the plastic tube. In both events, the object was lowered into the container and became invisible. That is, the short container event violated the ‘continuity’ principal of objects (i.e., an object placed inside a container should continue to exist without changing its shape) (Spelke, 1994). The authors reported that infants did not reliably show increased looking times for the short container event until around 7.5 months of age (Hespos & Baillargeon, 2001a). Furthermore, when a short or tall plastic tube was used as a cover, rather than as a container, it was only around 12 months of age that infants reliably look longer at an VoE event, suggesting their detection and processing of VoE (Wang et al., 2005). Taken together, it has been argued that infants perceive containment events differently from occlusion events, although all of these events are designed to test infants’ object permanence. Perhaps, occlusion events might be perceived by infants as more robust than container events to trigger their overt differential attentional behaviour. The discrepancy indicates infants’ perceptual processing systems during infancy takes time to develop to become more nuanced systems (Baillargeon & Wang, 2002; Luo & Baillargeon, 2005).

Given the quality of the current data, we would avoid making any assumption at this stage regarding how infants generally process unexpected events, or whether they do process various unexpected events in a general, cohesive manner. A limitation of the current study was likely to result from our attempt to use multiple video stimuli, while it was effective to ensure the sufficient amount of data necessary to conduct interpretable analyses. Although we did obtain enough number of analysable trials, the number of events included in the study might have been too many and different from one another to avoid latency jitter and data contamination. Nonetheless, we believe that the study has taken us one step closer to the goal of understanding what VoE means in the context of infant learning and cognitive development, by identifying a possibility of infants’ VoE processing being event-specific, and demonstrating a potential use of theta-band activity as a marker of infants’ processing of unexpected events.

Acknowledgement

I would like to thank all the infants and parents who took part in this study, and Dr Szilvia Linnert and Dr Katharina Kaduk for their invaluable advice on data acquisition as well as analysis. I would also extent my thanks to Barrie Usherwood for his immense help with creating stimuli and MATLAB coding. We also thank Marie Skłodowska-Curie Actions Innovative Training Network (MOTION) for their research funding and support.

Author contribution

Sayaka Kidby: design, stimuli making, data acquisition, analyses, writing, review. Kirsty Dunn: design, review. Vincent Reid: design, review.

Chapter 5

How infants learn novel words in a dynamic social interaction

Chapter Introduction

Research utilising electrophysiological measures to investigate infant cognition has advanced our understanding of the working of their cognition, even when their cognitive activity is not evident in their overt behaviour. Yet, a conventional approach of infant cognition research involving such measures typically utilised screen-based stimuli presentations, due to its proneness to gross motor movement. This means that evidence from these studies might overlook the complex interplay between infant cognition and social environmental factors, because such an experimental setting typically involves few or no social elements.

For instance, infant language learning has been investigated with neural measures, and these studies almost predominantly utilised a screen-based material (e.g., Friedrich & Friederici, 2008, 2011, 2017). However, for early language development, behavioural research consistently emphasises the importance of social components (Hakuno, Omori, Yamamoto, & Minagawa, 2017; Kuhl, 2007a; Kuhl, Tsao, & Liu, 2003). Such a disparity might perhaps contribute to the mixed evidence across studies regarding how infants might retain newly acquired lexical knowledge over time.

To address the issue, this chapter presents a study in which we investigated how infants encode semantic lexical information (i.e., words) during a naturalistic social interaction, and how acquired information leads to long-term memories, benefitting from offline consolidation. This provides an example of infant cognition research taking the second-person cognitive neuroscience approach, where infant brain activity was monitored whilst the experimenter and the child engaged in a structured, naturalistic social interaction.

This is a secondary data analysis. The original study was designed and conducted by Dr Katharina Kaduk for her doctoral thesis (2016). The aim of the original study was to investigate infants' retention of the knowledge of the association between novel nouns and objects presented during a social interaction. The retention was tested immediately after the learning event and on the following day. The data collected during the test phases were analysed focusing on Negative Central (Nc) and N400 ERP components, as well as theta and upper boundary of the alpha frequency band activity. Differences between the congruent and incongruent word-object pairs were found in the amplitude of the Nc, N400 and upper-alpha-band activity, indicating infants' discrimination of congruent and incongruent word-object pairs. Importantly, such differences were only observed in the data collected on the following day of the learning event, not immediately after. This supported the complementary learning systems framework, which stresses the importance of offline consolidation during which memory stored in the hippocampal system is transferred into the neocortical systems as long-term memory or 'knowledge' (McClelland, McNaughton, & O'Reilly, 1995).

As opposed to the analyses conducted by Kaduk (2016), which focused on the neural activity during the screen-based test phases, the current analysis focused on the neural activity during the learning event which took a live interaction format with the aim of exploring the online information encoding processes that led to offline consolidation as shown in the original analyses by Kaduk (2016). We further analysed how the neural activity during the learning event predicts infants' ability to discriminate congruent and incongruent word-object pairs in the subsequent two test phases conducted on two separate days. The current set of analyses presented in this thesis was hoped to address the question of how infants encode linguistic information in a social environment and how the underlying online neural process informs the subsequent word learning outcomes.

Abstract

For early language development, it is critical for infants to be able to appropriately associate an auditory label with a visual object. Research has suggested that infants can encode semantic (i.e., meanings) word-object associations with a very short presentation, but that their retention is rather weak (Friedrich & Friederici, 2008). Another body of research has suggested linguistic knowledge consolidation might only happen over time, for which sleep might play a critical role (Friedrich, Wilhelm, Born, & Friederici, 2015; Henderson, Weighall, & Gaskell, 2013). It is thus possible that infants can encode semantic information rapidly, but that consolidation may occur with a delay during which the information is stored as long-term memories. The current study investigated the trajectory of semantic word-object association learning in 10- and 11-month-old infants (*N*=20), using a combination of ERP and EEG frequency analyses. Infants were presented with novel words and novel toys in a live interaction, and the encoding and retention of the word-object semantic associations were assessed. Infants encoded semantic information during the live presentation as indexed by upper-alpha-band activity (9-10Hz), and evidence of semantic knowledge was only observed after a 24h delay, not immediately after the presentation. Further, individual differences were observed in terms of semantic information encoding and attentional engagement during a live social interaction, which, in turn, affected the degree of knowledge consolidation. The study provides evidence to support a complementary learning systems models of memory in the domain of infant word learning (McClelland et al., 1995), and demonstrates that individual differences in attention and information encoding in a social situation modulate knowledge consolidation processes.

Introduction

Language is important in every aspect of our social lives. It enables us to communicate, share ideas and transfer knowledge across cultures and generations. Accordingly, one of the important developmental achievements during infancy is to begin to learn words. It has been demonstrated that infants have the capacity to learn words to some extent even before they learn to speak. For instance, 6-month-old infants already associate the word “mama” and “papa” to their parents (Tincoff & Jusczyk, 1999). At around 8 months of age, infants begin to demonstrate the ability to extract novel words from fluent speech, based on their ability to detect patterns that frequently occur (Saffran, 2001; Saffran, Aslin, & Newport, 1996; Saffran & Kirkham, 2018).

The ability to form word-object associations with short exposure has been referred to as ‘fast-mapping’ (Carey, 2010). It has been shown that infants as young as 3 months of age have this ability (Friedrich & Friederici, 2017). In this study, 3-month-old infants were presented with 16 pseudo-words and 16 novel objects. Eight word-object pairs were consistently shown together, whilst the rest were randomly matched each time. Analysing event-related potentials (ERPs), the authors reported evidence that infants learned the word-object associations within only eight presentations, as reflected in the late-occurring negative component as well as the Pb component (Friedrich & Friederici, 2017). Yet, the N400 ERP component was not found, which has been associated with semantic (i.e., meanings) processing (Friedrich & Friederici, 2017). This indicates that infants aged 3 months are able to form word-object associations, but the level of understanding does not extend to semantic properties. However, using a similar experimental setup, it has been reported that older infants aged 6 months do show the N400 effect (Friedrich & Friederici, 2011). Thus, it is likely that the fast-mapping ability develops with age, and at the age of 6 months, this process starts to involve a lexical-semantic learning (Friedrich & Friederici, 2011, 2017) as opposed to lower-level perceptual associations.

Whilst learning the link between an object and its label is important, infants also need to integrate newly acquired linguistic knowledge into the existing semantic and lexical networks, so that the knowledge could be retrieved in the future (Weighall, Henderson, Barr, Cairney, & Gaskell, 2017). Such a ‘knowledge consolidation’ process has been proposed to be rather unstable during infancy (Friedrich & Friederici, 2008). For instance, infants younger than 14 months of age do show ERP components that suggest a certain degree of information encoding during the stimuli presentation (e.g., Pb and late negativity for 3-month-olds, N400 for 6-month-olds) but these components were not found the following day (Friedrich & Friederici, 2011, 2017). Evidence indicates that the development in knowledge consolidation occurs between 12 months and 14 months of age, as 14-month-olds show the N400 effect during the stimuli presentation as well as at the memory retention test taking place one day after the presentation (Friedrich & Friederici, 2008), whereas 12-month-olds do not (Friedrich & Friederici, 2010). When combined, these findings indicate that infants in their first postnatal year might struggle to consolidate information that has been encoded in a short time as long-term knowledge, and hence a stable memory structure for the encoded information might only last temporarily (no longer than one day).

On the other hand, it has been posited that word learning including lexical knowledge consolidation during infancy and early childhood is partial, slow and incremental (Brown, Weighall, Henderson, & Gaskell, 2012; Dockrell, Braisby, & Best, 2007; Henderson, Weighall, & Gaskell, 2013; Nagy & Scott, 216). The complementary learning systems of memory framework, or dual-memory systems theory, has proposed that novel information is initially stored separately from existing knowledge and gradually integrated over time ('offline consolidation') (Davis & Gaskell, 2009; McClelland et al., 1995; Norman & O'Reilly, 2003; O'Reilly & Norman, 2002; Weighall et al., 2017). Sleep is known to play an important role in the process of offline consolidation of newly learned words, as school-aged children show improved recall on the following day rather than immediately after the learning (Friedrich et al., 2015; Henderson et al., 2013; James, Gaskell, Weighall, & Henderson, 2017). Given the evidence with older children, Friedrich and colleagues (2015) compared memory test performance between infants who took a nap between an information presentation phase and a memory test phase and those who did not. The memory test took place 1.5 hours after the presentation phase. They found that infants who napped showed the N200-500 and N400 ERP components in the memory test phase, whereas those who did not nap showed no such effects. This suggests that sleep facilitated memory consolidation processes (Friedrich et al., 2015). Yet, sleep does not seem to fully explain other findings reporting infants' weak retention of knowledge mentioned above (Friedrich & Friederici, 2010, 2011, 2017). In these studies, infants' memory was assessed one day after the stimuli presentation session. Hence, these infants must have had sleep between the information encoding phase and the memory retention test (Friedrich & Friederici, 2010, 2011, 2017). Therefore, the evidence is currently mixed regarding the trajectory of infants' consolidation of semantic lexical knowledge. A possible explanation for the discrepancy might be that the absence of additional stimulation in Friedrich et al. (2015) between information encoding and retention test, as compared to other studies that were conducted over two days, might have played some role in the results that they obtained.

Mixed evidence might also be explained by the differences in experimental paradigms used. Typically, the neural process of infant word learning has been studied in a screen-based paradigm, isolated from social contexts (e.g., Friedrich & Friederici, 2010, 2011, 2017). However, a body of evidence suggests the critical importance of social settings in infant language learning (e.g., Kuhl, 2007). It has been shown that statistical learning of stimuli associations can occur from simple information input only, whereas the leaning of more complex and subtle information, such as speech sounds, needs multimodal input embedded in a social context (Kuhl, Tsao, & Liu, 2003; Maye, Werker, & Gerken, 2002; Saffran, Aslin, & Newport, 1996). Accordingly, evidence demonstrates that infants particularly benefit from live face-to-face interactions when learning language, compared to screen-based or audio-only interactions. Kuhl and colleagues (2003) had 9- and 10-month-old infants from English-speaking families interact with an experimenter in-person, or with a pre-recorded material presented on a screen. The experimenter spoke to an infant in Mandarin Chinese, and it was tested whether the infants had learned Chinese phonetics after the interaction. The results suggested that infants only learned the phonetics from a live interaction but not from a screen-based material. A similar effect of live interactions has been reported in terms of infants' ability to extract words from a fluent continuous speech (e.g., Hakuno, Omori, Yamamoto, & Minagawa, 2017). Hakuno and colleagues (2017) showed that infants who interacted with an adult in a live manner were successful in segmenting words from a spoken sentence, but those who interacted with an adult via screen did not (Hakuno et al., 2017). These behavioural studies indicate that a live contingent social environment facilitates infants' language learning. Therefore, it is possible that learning in a non-social context might not be as robust in terms of information encoding and the subsequent knowledge consolidation.

In addition to behavioural evidence, neural evidence has also suggested that infants' language learning is facilitated by social communicative cues (e.g., Conboy et al., 2015; Kuhl et al., 2008; Parise & Csibra, 2012; Rivera-Gaxiola, Silva-Pereyra, & Kuhl, 2005). For instance, Conboy and colleagues (2015) examined the effect of live social interactions on 10-month-old infants' learning phonetics of a foreign language. Analysing infants' looking behaviour and an ERP (mismatch negativity, or MMN), the authors found that joint attention with a foreign language speaker during an interaction phase predicted infants' successful discrimination between different foreign speech sounds at a subsequent test phase, as indexed by the MMN ERP component (Cheour et al., 1998; Conboy et al., 2015). This indicates that joint attention episodes during learning aid infants' learning of phonetics. Yet, whether this finding extends to other aspects of language learning, such as novel word learning, is not currently known. Another evidence demonstrating the role of social contexts on infant word learning was provided by Parise and Csibra (2012). The authors investigated whether 9-month-old infants hold semantic word-object associations, using ERP paradigms with a live interactional element. In this study, infants were seated in front of the monitor presenting video stimuli, with their mother and an experimenter seated on either side. Instead of pre-recorded auditory stimuli, either their mother or the experimenter uttered a name of familiar objects before each of the objects appeared on the screen. They found the N400 ERP component when the target word was uttered by the infant's mother, but no N400 was evident when the target word was delivered by an experimenter (Parise & Csibra, 2012). Whilst it is unclear exactly what factor drove this difference between a mother and an experimenter, this study demonstrates infants' word recognition may be modulated by social factors, such as social closeness or familiarity.

Taken together, being exposed to novel words in a social setting is critical for infants to successfully learn such complex entities as language, involving phonetics, phonology, syntax and semantics. It is hence plausible that infants might better encode lexical information deriving from a dynamic social situation as compared to in a controlled lab setting. Furthermore, it is possible that information encoding facilitated by social interactional elements might lead to a more robust memory structure that can last longer-term. The existing evidence with older children gives indirect support for this idea. In a study with 5- to 9-year-old children, their memory lasted longer when children were given cues to remember novel words, such as their meaning, as compared to when they were not (Henderson et al., 2013). This suggests that more robust information encoding supported by additional elements can lead to more robust long-term memory (Henderson et al., 2013).

Despite these findings suggesting the importance of social settings for infants to learn language, prior studies with neural measures investigating infants' semantic lexical knowledge consolidation typically utilised a limited number of stimuli presented together on a screen while controlling irrelevant information, such as distractor words which would be present in a naturalistic continuous speech. (e.g., Friedrich & Friederici, 2010, 2011, 2017). These experimental controls are effective to reduce the cognitive workload for infants. Yet, such an experimental setting where infants have little access to social information when learning novel words might not provide a very accurate picture of infants' language learning process that emerges in the real world. On the other hand, the existing studies investigating the role of social settings on infant word learning have not addressed the issue of how information learned in a social interactional setting might lead to a successful consolidation, as they have typically focused on the learning outcome assessed immediately after the learning session. Moreover, the online process of information encoding cannot directly be assessed using behavioural measures, and whilst possible, it has not been investigated using neural measures. Therefore, the neural and cognitive processes underlying the information encoding process within social interactions is still unclear. Furthermore, little is known about how information encoding occurs in a social interaction might influence the subsequent knowledge consolidation.

Currently, to our knowledge, there has been only one study using neural measures which presented infants with novel words and objects in a live interactional setting. Kaduk (2016) investigated the consolidation process of infants' knowledge of semantic word-object associations encoded during a social interaction. The author had 10- and 11-month-old infants interact with an experimenter in a live manner. During the interaction, the experimenter presented infants with two novel objects while repeatedly labelling the objects. Immediately after the interaction (without delay) as well as on the following day (with delay), infants were presented with stimuli showing congruent and incongruent word-object pairs. The congruent pair included the object appearing on a monitor after the 'correct' object label according to the live interaction on the previous day, and the incongruent pair included the object and the 'incorrect' label. They found differences between the conditions on the delayed test, in terms of the Negative Central (Nc) and N400 ERP components, as well as the degree of upper-alpha desynchronisation, which indicate infants' differing attention allocation, detection of semantic incongruency, and access to semantic systems, respectively (Kutas & Hillyard, 1980; Reid et al., 2009;

Reynolds & Richards, 2005). Critically, such differences were not observed on the immediate test without delay (Kaduk, 2016). These results suggest infants need 'offline consolidation' to integrate newly learned semantic knowledge into the existing lexical knowledge systems. This is consistent with the dual-memory systems account, suggesting that infants' knowledge consolidation is slow and incremental (Brown, Weighall, Henderson, & Gaskell, 2012; Dockrell, Braisby, & Best, 2007; Henderson, Weighall, & Gaskell, 2013; Nagy & Scott, 216). Furthermore, this work supports the idea that infants' knowledge consolidation benefits from a social interactional situation when encoding information, as 10-month old infants in this study were able to retain their newly acquired semantic lexical knowledge for at least a day after the learning event which occurred in a social interactional setting, despite previous work stating otherwise when information was presented in a non-social setting (Friedrich & Friederici, 2010).

Whilst Kaduk (2016) provided evidence that infants' word knowledge acquired in a social interaction successfully consolidates after a delay, the trajectory from learning to consolidation is yet to be fully understood. This is mainly because the process of information encoding at the time of taking part in the social interaction was not explored in Kaduk (2016). This additional component is critical, as there might be individual differences across infants in terms of encoding and consolidating semantic and lexical information. It has been proposed that a complex structure of word learning is underpinned by the development of cognitive and pre-linguistic communicative skills, such as working memory, rather than skills specialised for vocabulary acquisition (Bloom, 2004; Nazzi & Bertoncini, 2003; Woodward, Markman, & Fitzsimmons, 1994). This suggests that more general cognitive skills engaged in a social situation, such as attention, might affect infants' word learning, and individual differences in these skills might have contributed to the mixed evidence in the current literature. Yet, there have been no studies that have investigated what cognitive processes are engaged during the encoding process of word-label associations in a social interaction. Hence, how individual differences in such cognitive processes might influence the subsequent knowledge consolidation is yet to be fully understood.

The use of neural correlates can indicate infants' cognitive activity during social interactions, and hence can enable an investigation of infants' online word learning process that happens in a social setting. In adult studies, theta-band oscillatory power has been associated with word learning, whilst no studies thus far have found a link between infants' word learning and theta-band oscillatory power. For example, Klimesch and colleagues (1996) assessed whether theta-band activity can index successful word encoding in adults (Klimesch, Doppelmayr, Russegger, & Pachinger, 1996). In this study, participants were first asked to categorise a set of words into "living" or "non-living", and later they were asked to recall the words they categorised. A difference in theta oscillatory power was shown to successfully differentiate between recalled words and non-recalled words (Klimesch, Doppelmayr, Russegger, & Pachinger, 1996). This recall performance is likely to be incidental memory establishment, rather than active or intentional learning, due to the study design where the participants were not asked to remember the words prior to the recall task. Nevertheless, the relationship among theta-band oscillations, information encoding and memory establishment has been reported in other studies with adult participants (Guderian, Schott, Richardson-Klavehn, & Düzel, 2009; Weiss & Rappelsberger, 2000). Hence, it is possible that theta-band activity can index successful encoding of novel words, which leads to robust memory to be retrieved later.

In research with developmental populations, theta-band oscillatory activity has been associated with infants' learning in general in social and non-social contexts (Begus & Bonawitz, 2020). St. John and colleagues (2016) found differences in 12-month-old infants' brain activity across social and non-social contexts, suggesting that infants' cognitive processes are modulated by social interactional factors such as eye contact (St. John et al., 2016). Begus and colleagues (2015) reported that differences in theta-band oscillatory power during infants' object exploration was associated with differences in their looking time at a subsequent object recognition test (Begus, Southgate, & Gliga, 2015). This indicates that enhanced theta-band activity might index better object encoding that leads to more robust memory as compared to other objects. Not only *when* learning occurs, theta-band activity has also been found *before* learning occurs (Begus, Gliga, & Southgate, 2016). Another study conducted by Begus and colleagues (2016) assessed infants' theta-band activity when infants were faced with a speaker of their first language or a speaker of a foreign language. Given that infants prefer to look at in-group members, such as speakers of the same language, they expected to observe increased theta-power activity towards a native speaker. The results supported their hypothesis, and suggested a potential link between theta-band oscillations and infants' motivation to learn from the speaker, expecting to receive information from an in-group member (Begus et al., 2016; Csibra and Gergely, 2009). The role of theta-band activities in infants' expectation towards other people has also been reported in a peek-a-boo game, where infants can expect their play partner to appear, disappear and re-appear (Orekhova, Stroganova, & Posikera, 1999). These studies propose an association between theta-band activity and infants' anticipatory attention. Anticipatory attention is relevant to the context of infants' word learning in a social interactional situation, as social interactions are thought to prepare infants to learn from others so they expect other people to provide them relevant information to learn (Csibra & Gergely, 2009). Other studies have associated infants' theta-band activity with attention more generally (Bosseler et al., 2013). For instance, during a naturalistic play session with a parent, theta-band activity in 12-month old infants has been found to precede sustained attention, and to be positively correlated with the duration of attention fixation (Wass, Clackson, et al., 2018). Taken together, the examination of theta-band activity might clarify cognitive processes underlying infants' word learning in a social interactional setting.

For infant language learning, alpha-band activity might also be relevant. It has been associated with attention as well as semantic processing and memory, with the latter particularly being reflected in the upper boundary of the alpha band (Başar & Güntekin, 2012; Klimesch, 2012; Neuper & Pfurtscheller, 2001). Alpha-band activity has been thought to represent a voluntary and controlled allocation of attention (Neuper & Pfurtscheller, 2001). In infant research, studies that used a gaze cueing paradigm and conducted event-related oscillation analyses (or a time-frequency analysis) have also interpreted alpha-band activity in terms of attention (e.g., Hoehl, Michel, Reid, Parise, & Striano, 2014; Michel et al., 2015). Hoehl and colleagues (2014) monitored 8- and 9-month-old infants' brain activity whilst they looked at an object displayed on a screen. There was an experimenter in the same room as the infant. In one condition ('Joint Attention' condition), the experimenter looked at both the infant and a screen showing an object in an alternate manner. In another condition ('Control'), the experimenter looked only at the screen, avoiding eye contact with the infant. The authors found desynchronised alpha-band activity in the Joint Attention condition but not in the Control, and argued that infants' alpha desynchronisation was induced by social triadic interactions scaffolding infants' attention (Hoehl, Michel, et al., 2014). Michel and colleagues (2015) presented 2-, 4-, 5-, and 9-month-old infants with a set of photos showing an object and an adult. Half of the photos showed an adult looking towards the object ('object-directed'), whereas the other half of the photos showed an adult looking away from the object ('object-averted'). They found alpha desynchronisation in 4-month-olds and 9-month-olds while they were presented with the object-directed gaze photos. Consistent with Hoehl et al. (2014), these alpha desynchronisation effects were interpreted as infants' increased attention towards social cues (i.e., eye gaze) (Michel et al., 2015). Given these reports, it is plausible that alpha desynchronisation reflects infants' attention towards social information, such as other people giving cues to attentional target, whilst such increased attention might be a response to external perceptual stimuli rather than infants' voluntary attentional control (Wass, Clackson, et al., 2018). Nevertheless, alpha-band desynchronisation could be an index of infants' attentional and cognitive processes underlying infants' word learning in a social interaction.

Relatedly, the upper bound of alpha frequency band has been found in relation to semantic processing in adults (Klimesch et al., 1996; Klimesch, 2012) and infants (Kaduk, 2016). Klimesch and colleagues (1996) presented adult participants with 96 object-concept pairs (e.g., "claws-eagle", "bananas-yellow") and asked them to evaluate them as semantically congruent or incongruent. Following this task, the authors asked the participants to report as many related concepts as possible to a certain object (e.g., "eagle"), while the brain activity of the participant was monitored. The authors found that upper-alpha desynchronisation effect was stronger in participants who were able to name more concepts than the average ('good performers') than in those who named less than the average ('bad performers'). The authors also assessed how many concepts the participants remembered from the previous presentation of the 96 object-concept pairs, and this simple recall performance did not explain the difference between 'good' and 'bad' performers. Given the result, it was argued that upper-alpha desynchronisation should reflect a semantic understanding or concept of the object, rather than a mere memory trace of the previous task (Klimesch et al., 1996). With infant participants, Kaduk (2016) examined infants' recall of newly learned words via upper-alpha-band activity. They showed that the different degree of infants' recall was indexed by the presence of absence of N400, and importantly, also by the upper-alpha-band activity. In infancy research, the N400 has been identified as a reliable neural marker for semantic processing in various domains including language (e.g., Friedrich & Friederici, 2006; Kutas & Hillyard, 1980). Their correlation analysis that found the association between N400 and upper-alpha-band activity suggests that the upper-alpha-band activity can be reliably used as an index of semantic processing in infants (Kaduk, 2016). The importance of this finding should be emphasised given that studies utilising N400 have produced mixed evidence as to whether infants can retain newly acquired knowledge overnight (e.g., Friedrich & Friederici, 2005, 2011; Friedrich et al., 2015). Adding another neural marker to investigate could increase the sensitivity of measurement, enabling a detection of another aspect of semantic knowledge consolidation.

The current study aims to explore the dynamics of infant word learning occurring in a naturalistic social interaction, as reflected in neural oscillatory activity in theta and alpha bands. Furthermore, it also aims to examine the association between the neural activity during the information encoding phase, which occurs taking a form of social interaction, and during the memory retention tests on the same day and the following day. The study is hoped to address some of the unanswered questions regarding how infants encode novel word-object associations in a dynamic and complex social setting, and how individual differences in the process of information encoding, if any, might influence the subsequent knowledge consolidation process. As it has been proposed in adult studies that alpha band should not be considered as a unitary band (Klimesch, 2012), the current study examines infant theta-band activity (3-5Hz), alpha-band activity (6-9Hz) and upper-alpha-band activity (9-10Hz) separately, as an index of infants' encoding of novel information, attentional allocation and semantic understanding, respectively. Based on prior studies, we hypothesise that infants' theta oscillatory power increases and alpha oscillatory power decreases (i.e., becomes more suppressed) as infants hear more instances of object labelling uttered by an experimenter in a live interaction, given that increased theta oscillatory activity has been linked to infants' increased attention and information encoding (Begus et al., 2016, 2015; Wass, Noreika, et al., 2018), and supressed alpha-band activity to attentional allocation as well as the formation of knowledge and memory (Hoehl, Michel, Reid, Parise, & Striano, 2014; Klimesch, 2012; Michel et al., 2015). Furthermore, we expect the activity in the upper-alpha frequency band to be more supressed as infants hear more labelling, as the supressed upper-alpha-band activity has been linked to better knowledge consolidation (Klimesch, 2012, Kaduk, 2016).

The current study analyses infant EEG data collected during an interaction with an adult experimenter who presented novel objects and labelled them with pseudo-words. This learning phase was followed by two recognition test phases, conducted on the same day and the following day. The data collected during the recognition test phases were analysed by Kaduk (2016) using ERP and ERO paradigms. Our analysis focuses on the learning phase, and the link between the data from the learning phase and the data during the two memory test phases. The results of the memory test phases were provided by Kaduk (2016) to be included in the current study. The analysis of the association between the learning and test phases enables the evaluation of the trajectory of knowledge consolidation. We expect that infants who showed a stronger learning effect during the learning phase involving social interactions would also exhibit stronger evidence for knowledge consolidation with a delay after the learning phase, benefiting from robust information encoding and offline consolidation (Friedrich et al., 2015; Kaduk, 2016). On the other hand, infants who showed a lower engagement in a learning interaction phase are expected to show a weaker retention even after a delay, reflecting individual differences in social cognitive skills impacting information encoding in a social situation as well as knowledge consolidation processes (Friedrich & Friederici, 2010).

Methods

Participants

The final sample consisted of 20 infants (12 females) aged between 10 and 11 months (M_{age}= 322.75 days ranging between 248 and 347 days, SD_{age}= 22.37 days). All infants were born full-term (37-42 weeks) in the normal range for birthweight (over 2500 grams). The primary language of the participating families was English. Additional 15 infants were tested but excluded due to fussiness (*n*=3), and an insufficient number of usable EEG data necessary for data averaging (i.e., less than five) (*n*=12).

The initial analysis comparing three blocks of the learning phase was conducted on the data collected from 15 infants (8 females) aged between 10 and 11 months (M_{age}= 320.87 days ranging from 248 to 347 days, SD_{age}= 24.33 days). The sample for this initial analysis was composed of infants who contributed to at least five usable epochs to all the three blocks of the learning phase. As described below, however, based on the behavioural data, the third block of the learning phase was excluded from the final analysis. Hence, the final sample consisted of infants who contributed at least five usable trials to both the first and second blocks of the learning phase, but not necessarily to the third block. This led to an increase in sample size for the final analyses.

Data acquisition

The current study is a secondary data analysis of the data collected by Dr Kirsty Dunn and Dr Katharina Kaduk for another study conducted in 2014 and 2015, which explored infants' memory consolidation of semantic information using event-related potentials (ERPs) and event-related oscillations (EROs). All infants were recruited through phone calls and emails from a database consisting of caregivers who expressed an interest in taking part in psychological research. The original study involved two separate visits to Lancaster Babylab, for which the caregivers received travel imbursement and a book as a gift for participation. Prior to recruitment, ethical approval regarding the study procedure, including research methods and data handling throughout the study, was sought and obtained from the Lancaster University Faculty of Science and Technology Research Ethics Committee. All the parents whose child participated in the current study were given a full description about the study from the experimenter, and gave a written informed consent on behalf of their child.

Stimuli

Two novel objects (Object A and B, see Figure 5-1) were selected as stimuli to present. The dimensional properties of the objects were similar with negligible differences (Object A: 11.4 x 11.4 x 3.2 cm, Object B: 12.1 x 10.2 x 4.4 cm). Objects were labelled with pseudo-words ("YOK" and "BLAP"). The object-label pair as well as the presentation order were counterbalanced across participants, but consistent for each participant (i.e., the same label was given to the same object throughout the learning phase for the same participant).

Figure 5-1. Illustrations of Objects A (left) and B (right) presented to infants. This figure is reused from Kaduk (2016) with permission.

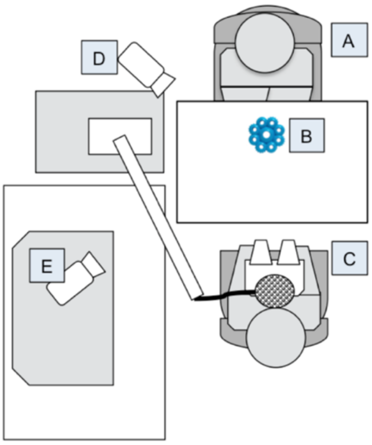
Procedure and apparatus

Infants were seated on the parent’s lap facing a table. An experimenter was seated at the opposite side of the table. Two camcorders were placed to videorecord behaviour of the infant and the experimenter. This was to enable offline coding for a behavioural analysis (see the relevant section below).

Infants underwent the learning phase, and the two recognition test phases. This study focused on the data collected during the learning phase, during which infants were presented with two novel objects, one at a time, by an experimenter in a live interaction. This learning phase consisted of six sessions in total, in which each object was presented three times in an alternate manner. Each session lasted for one minute.

At the beginning of each session, the experimenter took an object out of the box that was placed under the table, and demonstrated its function to the infant. Each session started with the experimenter holding up the object and saying, “*Look what I have here!*” and labelling the object as “*Look! It’s a [YOK/BLAP]!*”. During the demonstration, the experimenter interacted with an infant in a naturalistic manner, such as calling the infant by their name and maintaining direct eye contact. The experimenter labelled the object as often as possible to facilitate the infant’s information encoding. On average, she labelled an object 17.3 times during a 1-minute session (*SD* = 1.3).

Figure 5-2. Illustrations of the experimental setup for the data collection.

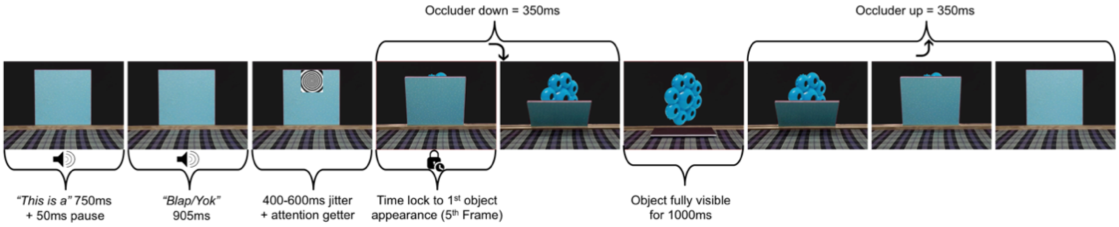


Note. An experimenter (A) sat at the opposite side of the table to the infant (C), and presented them with an object (B). The infant (C) sat on their parent’s lap. The entire session was video-recorded by two camcorders (D, E), recording behaviour of the experimenter and the infant. This figure is reused from Kaduk (2016) with permission.

Two recognition test sessions followed the learning phase. The first test session ('Day 1 test phase') took place immediately after the learning phase on the same day. Infants and their caregivers faced the CRT 19-inch monitor displaying 3D stimuli, placed 90cm apart from the participants. The second test phase ('Day 2 test phase') was conducted the following day. The procedure was identical to the first test phase.

The stimuli for the test phases consisted of a series of images showing one of the objects presented during the learning phase, as well as an occluder. Each sequence started with a female voice saying, "*this is a…". The* voice was not of the experimenter who interacted with the infant during the learning phase. This was followed by a 50ms pause, after which the word prime "YOK" or "BLAP" (905ms long each) was introduced by the same female voice via a speaker. An attention getter appeared on the top half of the occluder for a random duration of 400-600ms before the occluder slowly started to be lowered to show one of the objects from the learning phase. The object remained visible for 1000ms. After this, the occluder moved upwards for 350ms to occlude the object again. The next sequence started after a 400ms-long interval. The half of the sequences presented a matching pair of object and label from the learning phase ('congruent condition'), while the other half of the sequences presented an unmatched pair ('incongruent condition'). When the infant became inattentive, the experimenter introduced an attention getter on the monitor, or temporarily paused the stimuli so the infant could take a break. Each test phase lasted until the infant could no longer sustain attention to the stimuli.

Figure 5-3. Illustration of the stimuli presented during the recognition phase.



Note. Each sequence started with a female voice saying "this is a…" (750ms), which was followed by a 50ms pause. Then, "YOK or "BLAP" was introduced by the same female voice, after which the attention getter appeared on the top half of the occluder. The object was presented as the occluder was lowered, and remained fully visible for 1000ms. The occluder was then moved upwards for 350ms to fully cover the object again. The word-object pair was either congruent or incongruent based on the paring introduced during the learning phase. The figure is reused from Kaduk (2016) with permission.

EEG recording

EEG data were collected using a 128-channel Geodesic Sensor Net (HCGSN 130; EGI, Eugene, OR). EEG data was amplified with an EGI Net Amps 300 amplifier with a sampling rate of 500Hz. (HCGSN 130, EGI, Eugene, OR, USA). EEG was online referenced to the vertex electrode (Cz). On storing the data for subsequent offline analyses, an analogue bandpass filter (0.1 to 100 Hz) was applied. The four channels placed on the forehead (channels 125, 126, 127, 128) were disconnected to avoid excessive discomfort for infant participants of having electrodes close to their eyes. Thus, the data for the current study were collected from 124 channels in total.

Behavioural analysis

Video recordings from the learning phase were coded frame by frame using ELAN 5.9 to mark time intervals during which the experimenter labelled the object. To ensure the accuracy of the coding, 25% of the data were coded by a secondary coder. On average, the difference between the primary and the secondary coders was 0.16 seconds (*SD* = 0.091 sec). As the current study only extracted the power spectrum over 2-second epochs, this difference (0.008% relative to the duration of an epoch of focus) was considered to be negligible enough to not influence the results.

Time intervals during which the infant did not visually attend to the experimenter or the object were also coded using the same video recording. These time intervals were excluded from the analysis. Whilst the focus of the current study was to examine the neural marker of infants' word learning from auditory input, it was considered to be most appropriate to use observable behavioural cues (i.e., whether they were looking at the experimenter or the object) to assess their attention to auditory information, as it is challenging to precisely detect infants' attentional level to auditory stimuli from their overt behaviour. For an exploratory purpose, time intervals during which the infant had physical contact with the object but not looked at the experimenter or the object (e.g., the infant was holding the object, not just coincidentally resting their hand on the object, while looking elsewhere) were also coded. This was done to increase the number of segments included in the analysis, and hence to improve the signal-to-noise ratio. These intervals were considered to be relevant because voluntary physical contact suggests infants' engagement with the object, which might lead to the infant associating the object with the utterance of the object label even with the absence of overt visual attentional behaviour. Nevertheless, we did obtain a sufficient number of usable segments without including these time intervals. Therefore, the analysis only included the time intervals during which infants showed visual attentional engagement to the experimenter or the object. To validate the behavioural coding, 25% of the videos were coded by a second coder separately for attention and the timing of the labelling. The initial percent agreement rate was 93.16%, Cohen's $\kappa = .90$. The trials for which the coding did not match between the coders were reviewed and recoded by the two coders.

Table 5-1. Coding scheme used for attention coding.

Coding scheme	
1	Looking: Looking at the experimenter or the object. - When the object is moving, the infant head movement follows the trajectory. - When the object is on the table, the infant' gaze should be directed to the toy. When unclear, code 1 if the baby is looking at the correct side (right, centre or left) of the table (i.e., where the toy is placed). - When the object is held by the experimenter above the table (and it is difficult to decide what the infant is looking at), code 1 if the baby is looking at the correct side (right, centre or left) relative to the experimenter.
0	Not looking: Infants are not looking at either the object or the experimenter.
0.1	Not looking but touching: infants are not looking at the experimenter or the object, but their hand has contact with the object.

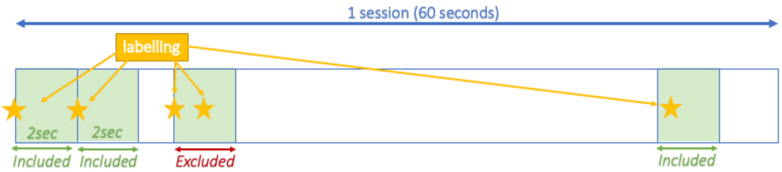
EEG pre-processing

EEG data collected during the learning phase were segmented according to the onset of each session using the Net Station 4.5.4. The segmented data (6 x 1-minute sessions, concatenated) were then exported as a MATLAB compatible format and imported to EEGLAB v2021.0 on MATLAB 2019b. The data were further filtered between 1 and 12 Hz. At the time of pre-registration, this filtering range was set to 1 and 50Hz. However, visual inspection of the data suggested a high level of noise at high frequencies. Since the focus of the analysis in the current study was alpha- and theta-bands (3-9Hz), the decision was made to filter out the signals of higher frequency than 12Hz to enable manual artefact detection.

The labelling timing codes were also imported and attached to each dataset as 'events'. Based on these event codes, 2-second-long epochs were extracted, so each epoch would start with the timing when the experimenter finished saying an object label (see Figure 5-4). Epochs containing more than one labelling were not included in the analysis, as they might include neural responses to auditory stimuli.

The epoched data were then visually screened for artefacts. Channels that were contaminated from eye movements, eye blinks, and gross motor movements were marked as noisy. Epochs containing more than 10 channels with excessive noise levels were discarded. For the remaining epochs, channels which were marked as noisy were interpolated by using the average of surrounding electrodes. After the artefact detection, the data were re-referenced to the average electrode using the EEGLAB average re-reference function (pop_reref.m).

Figure 5-4. Illustration of how epochs were extracted.



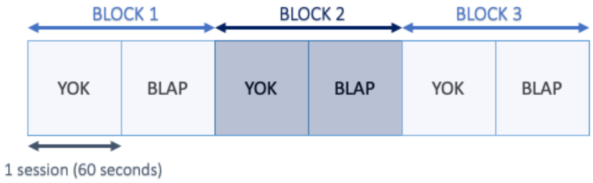
Note. Each epoch was segmented according to the behavioural coding data, which marked the offset of the object labelling event (i.e., when the experimenter finished saying an object label). Each epoch was two-second long. Epochs that contained more than two labelling were excluded to avoid confounding with a neural response to auditory stimuli.

EEG analysis

Learning phase

To examine the theta- and alpha-band neural activity as an index of infants' novel word learning over the course of the learning phase, the first two sessions, the middle two sessions and the last two sessions of the learning phase were grouped together, and the resulting three blocks were examined in the current study (see Figure 5-5). Only infants who contributed more than five usable epochs to each block were included in the final analysis. The threshold was set based on the previous studies that reported interpretable results on the data from infants who contributed as few as three to seven trials per condition on a standard event-related potential (ERP) paradigm (Kaduk et al., 2013; Stets & Reid, 2011).

Figure 5-5. Illustration of the data structure used in the current study.



Note. The data collected during early, middle, and last two sessions were grouped together, making three blocks in total. The current study explored the difference across the three blocks.

The 2-second artefact-free epochs were submitted to fast Fourier transformations (FFTs), using the EEGLAB spectopo.m function which performed FFTs with Hamming-tapered windowing of 500ms with 50% overlap (250ms). Power estimates were extracted between frequencies between 1 and 12Hz in steps of 1Hz.

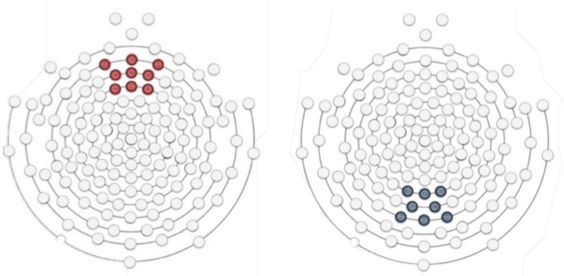
Grand averages of the FFTs were computed for each included infant for each block, and three estimates of power spectrum were yielded for each participant. For the statistical tests, the $\log_{10}$ transformation was performed over the data to improve the normality of the amplitude distribution. Power estimates from each participant were averaged for each block, and these were compared across blocks to examine the temporal dynamics of neural activity while infants repeatedly heard the experimenter label the object and form word-object associations in a live interaction.

Based on previous studies, theta-band oscillatory activity (3-5Hz) over the fronto-central electrode sites (Ch4, 9, 10, 11, 15, 16,18, 19, 21) was analysed to investigate differences in neural activity, as it has been associated with information processing and learning involving social interactional factors (Begus, Southgate, & Gliga, 2015a; Hoehl, Michel, Reid, Parise, & Striano, 2014; Michel et al., 2015; Orekhova, Stroganova, Posikera, & Elam, 2006; Saby & Marshall, 2012). At the time of pre-registration (https://osf.io/p5bq8/), the theta-band range was defined as 3 to 6 Hz. However, to make the analysis comparable to Kaduk's (2016), the frequency band was re-defined to 3-5Hz. 3-5Hz has also been frequently used as the range of infant theta in other studies (Begus et al., 2016, 2015).

Alpha-band activity over posterior occipital sites was also examined as it has been linked to voluntary attentional processes to social stimuli (Orekhova, Stroganova, & Posikera, 2001). Alpha frequency band in infant populations is typically defined as from 6 to 9Hz

(Cuevas, Cannon, Yoo, & Fox, 2014). Therefore, the current study analysed this frequency range. The previous study by Kaduk (2016) analysed the data from the same sample and found an effect over activity in the frequency range of 9 to 10Hz over eight parietal-occipital channels (Ch67, 70, 71, 72, 75, 76, 77, 83). Hence, this upper alpha band over the same parietal-occipital area was also examined. The electrodes used for each analysis were shown in Figure 5-6.

Figure 5-6. Channels used to analyse oscillatory activity in theta (left), alpha and upper-alpha band (right).



Note. Fronto-central nine channels (Ch4, 9, 10, 11, 15, 16,18, 19, 21) were chosen as a focus of analysis on theta-band activity, and occipital-parietal eight channels (Ch67, 70, 71, 72, 75, 76, 77, 83) were chosen for the analysis of alpha-band and upper-alpha-band activity.

Association between the learning phase and the test phases

The relationship between the neural activity during the learning phase (theta-, alpha- and upper-alpha-band oscillatory power) and the neural activity during the test phases on Days 1 and 2 (event-related oscillatory power in theta- and upper-alpha-band, as well as the Nc and N400 ERP components) was assessed using linear mixed-effects models (LMMs). LMMs can account for the nesting structure of the data by considering between-participant variability. As the current data included repeated measures (theta-, upper-alpha, Nc and N400 measured on the test phases on Days 1 and 2), LMMs were considered to be more optimal than multiple linear regression models. Due to the within-subject design of the analysis, we included infants who met the inclusion criteria for both the current study (i.e., five or more usable trials per block) and the study conducted by Kaduk (2016) (i.e., four or more usable trials per condition). The resulting sample size was *n*=14.

The aim of this analysis using LMMs is twofold: First, we aimed to evaluate how the neural dynamics observed during the learning phase which took place on Day 1 might predict infant word learning outcomes reflected in differences in neural activity between the congruent and incongruent conditions (i.e., presentations of correct and incorrect word-object pairs) of the test phases which took place on Day 1 (immediately after the learning phase) and on Day 2 (24 hours after the learning phase). If any particular pattern of neural frequency activity across different blocks of the learning phase which took place on Day1 significantly predicts infants' successful discrimination of correct and incorrect word-object pairs on Days 1 and 2 test phases, which would be reflected in a significant difference in ERPs and frequency power between congruent and congruent conditions on the test phases on Days 1 and 2, it would enable us to discuss what cognitive abilities at play during the learning event informed the subsequent learning outcomes. Furthermore, this could propose neural activities in these frequency bands as valid measures of infant word learning occurring during a naturalistic social interaction. The analyses of ERPs and frequency power on test phases was reported in Kaduk (2016), of which results indicated that infants only show differential neural activity (Nc, N400 ERP components as well as upper-alpha-band oscillatory power) between congruent and incongruent conditions on Day 2 test, but not on Day 1 test. Given this piece of evidence, we expected that neural dynamics observed during the learning phase should predict the neural activity patterns that indicates infant word learning (i.e., differential Nc, N400, upper-alpha-band power between the congruent and incongruent conditions of the test) on the Day 2 test but not on Day 1. This would be reflected in the effect of test date in LMMs (the main effect or interaction). If the main effect of test date should be found, this would be a replication of Kaduk's (2016) analysis which should suggest that neural activity differences between congruent and incongruent conditions were only evident in Day 2 not in Day 1. If any interaction should be found, this would indicate that specific dynamics in certain frequency band (i.e., theta-, alpha-, and/or upper-alpha-bands) which occurred during the learning phase predict successful word learning reflected in neural activity differences between the congruent and incongruent conditions of the test phase, but the direction and/or degree of such predictive relationship differed between Days 1 and 2 (i.e., immediate or delayed test after the learning phase).

Secondly, through these LMM analyses, we aimed to investigate potential individual differences in neural dynamics underlying infant word learning in a social situation which affect their learning outcome. This was motivated by two previous reports: First, Kaduk (2016) found no significant difference in theta-band oscillatory power between the congruent and incongruent condition despite a growing evidence suggesting the use of theta-band power as an index of infant learning or information encoding (Begus & Bonawitz, 2020). This null result was interpreted that the learning which can be reflected in theta-band power must have happened during the learning phase on Day 1 before the test phases on Days 1 and 2 took place (Kaduk, 2016). Whilst the direct examination of this interpretation would be done by the analysis described focusing on neural dynamics in theta frequency band during the learning phase above (see ‘Learning phase’ of this EEG analysis section), the analyses using LMMs linking the neural activity of learning phase and test phases would also be informative in further investigating this possibility. That is, if theta-band power successfully indexes infants’ information encoding, the dynamics of theta-band activity during the learning phase (Day1) should predict infants’ discrimination of congruent and incongruent word-object pairs on the test phases (Days 1 and 2) as reflected in ERPs and upper-alpha-band activity as reported in Kaduk (2016). This would be assessed by the examination of the main effect of theta-band dynamics during the learning phase (Day 1) in LMMs. Secondly, as discussed above, there has been mixed evidence as to whether infants younger than 12 months of age can hold semantic knowledge overnight (e.g., Friedrich & Friederici, 2010, 2011, 2017; Kaduk, 2016). Potentially, this could well be due to individual differences in more domain-general, pre-linguistic cognitive skills, such as attention, which are at play during the learning event (Bloom, 2004; Nazzi & Bertoncini, 2003; Woodward et al., 1994). If this was the case, the neural dynamics of alpha-band oscillation during the learning phase, as an index of infant attention (Hoehl, Michel, et al., 2014; Michel et al., 2015) might well predict infants’ discrimination between congruent and incongruent conditions (i.e., correct and incorrect object-label pairs) at the test phases (Days 1 and 2). If such a relationship was found, this could explain the existing mixed results in the existing literature regarding why some studies fail to report infants’ successful retention of knowledge overnight (Friedrich & Friederici, 2010, 2011, 2017). Critically, given that social situations can facilitate infants’ sustained attention (Wass, Clackson, et al., 2018), this analysis could lead to a further discussion on the proposal that the lack of evidence that infants can hold memory for a longer term is due to the (non-)social nature of the learning event in an experimental setting, as well as that infants’ memory establishment and retention is more robust when information is presented in a social situation (Kuhl, 2007).

For the predictor variables of the LMMs, we calculated the difference between Blocks 1 and 2 during the learning phase in theta-, alpha- and upper-alpha frequency bands. The ‘difference scores’ were used as predictors in each LMM. Prior to the analysis, these predictor variables were re-scaled so the mean value would be 0 with the standard deviation being 1. Therefore, a negative value of this score would indicate the power in the frequency band of focus was greater during Block 1 than Block 2, whilst a positive value would suggest the power in the frequency band of focus was greater during Block 2 than Block 1.

The outcome variables were the difference between the congruent and incongruent conditions on the test phase on Days 1 and 2. Such ‘difference scores’ were calculated based on the data of the test phases obtained with permission from Kaduk (2016), following the method used in the regression analysis reported in her study. Specifically, for theta- and upper-alpha-band power, the mean power in the time window of focus was used to obtain the difference score. For the Nc, the mean power in the time window of focus was used. For the N400, the difference score was obtained by computing the difference in the area of ‘response curve’ (Hoormann, Falkenstein, Schwarzenau, & Hohnsbein, 1998; Kaduk, 2016). The straight line was placed on the EEG waveform in a way that it connects two points at either side of the time window of interest (650-800ms post stimulus onset), and the distance between the curve and the line at each sampled time (75 samples in total as per the sampling rate of 500Hz) was calculated. This resulted in a set of values representing a deviation of the curve from the straight line. The values were summed to create one score for each condition. A larger value of this score would indicate a larger deflection in the negative direction (i.e., N400) and a smaller value would suggest an absence of the N400 component (Kaduk, 2016). The calculations and interpretations are summarised in Table 5-2 and 5-3.

Bayesian LMMs were constructed using brms (Bürkner, 2017, 2018) and bayestestR (Makowski, Ben-Shachar, & Lüdecke, 2019) packages in R (R Core Team, 2014). The Bayesian models were considered optimal given the small sample size included in the analysis (*n*=14).

Table 5-2. Learning phase difference scores used in the linear mixed-effects model as predictors.

Learning phase difference score		
Frequency band	Calculation	Interpretation
Theta-band oscillatory power	M_{Block2} - M_{Block1}	*Positive* value indicates a greater power during Block 2 than Block 1
Alpha-band oscillatory power	M_{Block2} - M_{Block1}	*Positive* value indicates a greater power during Block 2 than Block 1
Upper-alpha-band oscillatory power	M_{Block2} - M_{Block1}	*Positive* value indicates a greater power during Block 2 than Block 1,

Table 5-3. Testing phase difference scores used in the linear mixed-effects model as response variables.

Test phase difference score		
Frequency band	Calculation	Interpretation
Evoked theta-band oscillatory power	$M_{congruent}$ - $M_{incongruent}$	*Positive* value indicates that infants showed greater power in the congruent condition.
Evoked upper-alpha-band oscillatory power	$M_{congruent}$ - $M_{incongruent}$	*Positive* value indicates that greater alpha-band power was observed in the congruent condition than the incongruent condition.
Nc ERP component	$M_{incongruent}$ - $M_{congruent}$	*Positive* value indicates that the mean amplitude of the Nc was greater in the congruent condition. *Negative* value indicates that the mean amplitude of the Nc was greater in the incongruent condition.
N400 ERP component	$M_{incongruent}$ - $M_{congruent}$	*Positive* value indicates that the N400 component was more prominent in the incongruent condition compared to the congruent condition.

Results

Behavioural analysis

Infants' visual attention was examined using the video recordings taken during the EEG data acquisition sessions. Infants who were too fussy to engage in the experimental task (n=3) were excluded from the behavioural coding, making the sample size for this analysis n=32.

A visual inspection of the video recordings suggested that infants' attention noticeably reduced during the Block 3, as compared to Blocks 1 and 2 of the learning phase. A repeated-measures ANOVA comparing the number of epochs (i.e., 2-second time windows after the experimenter's labelling) during which infants visually attended to the task taking Block (Blocks 1, 2, 3) and Label (YOK, BLAP) as within-subject measure found a main effect of Block ($F(2,61) = 19.41$, $p< .001$), with a pairwise post hoc test indicating that the number of epochs during which infants attended to the task was larger during Block 1 than Block 2 (p=.021), and also than Block 3 (p=.0001). Whilst there was no significant difference between Blocks 2 and 3 (p=.029) in the number of epochs during infants were attentive to the task, this null result may well be due to a large individual difference (i.e., standard deviation) (see Table 5-4).

Table 5-4. The number of epochs during which infants visually attended to the task

	N	M_{trials}	SD_{trials}
Block 1	32	16.05	2.91
Block 2	32	14.30	3.87
Block 3	32	13.64	4.16

Visual inspection of the EEG data also suggested the data quality markedly declined from Block1 to Block 2, and then to Block 3 in such a way that the data from Block 2 as well as from Block 3 involved a higher level of noise as compared to the data from Block 1. To further investigate the data quality, we examined the mean and standard deviation of artefact-free EEG epochs for each trial. This showed that a number of infants in the current sample contributed 0 usable epoch for Block 3 (see Table 5-5). Therefore, the data from Block 3 would only be from a very selected sample of infants. This would sacrifice the interpretability and generalisability of the analysis results, meaning that it would be necessary to exclude the data from Block 3 in the final EEG analysis to conduct meaningful and credible analyses.

Additionally, we statistically compared across Blocks in terms of the number of the number of artefact-free EEG epochs, via a repeated-measures ANOVA taking Block as a within-subject measure, given that the previous ANOVA did not find a main effect of Label. The descriptive statistics are shown in Table 5-5. The analysis found a significant decline in the number of artefact-free EEG trials between Blocks 1 and 3 (pairwise post-hoc test: p = .001), whilst no other differences were found. This also supports our decision that the quality of the data from Block 3 was not comparable to Block 1 during which infants were likely to be most attentive, contributing to a sufficient number of artefact-free data necessary for data averaging as part of reliable EEG analyses. Consistently, the statistical analysis examining differences in the EEG data attrition rate (i.e., the percentage of the EEG data excluded based on the level of noise and artefact) also suggested a significant difference between Blocks 1 and 2, and Blocks 1 and 3. That is, a repeated-measures ANOVA comparing 3 blocks, taking Block as a within-subject measure, found a significant main effect of Block (F(2,56) = 17, p <.001) (see Table 5-6 for descriptive statistics). A pair-wise post-hoc comparisons indicated higher attrition in infants' attention during Block 2 as compared to Block 1 (p = .002), and during Block 3 as compared to Block 1 (p = .001).

As the EEG attrition rate represents the portion of the data which were contaminated by motor or eye artefact (see EEG pre-processing above), this can be used as an index of infants' task engagement. Hence, the significant differences identified in the analysis of EEG attrition rate in the current sample is consistent with our interpretation of behavioural data suggesting a significant attenuation of infants' attention from Blocks 1 to 2, and to 3. Attenuation of infants' attentional engagement towards the end of experimental sessions has previously been reported, supporting this interpretation (Stets et al., 2013; Stets & Reid, 2011).

Infants' attention attenuation can also be understood as habituation reflecting the full encoding of information, or learning (Colombo & Mitchell, 2009; Nelson, 1995; Turk-Browne et al., 2006), whilst the decline in attention could also be a sign of fatigue (de Haan, 2013). Given that infants' attention did not recover between Blocks 2 and 3 when the presented object-label pair switches, it is likely that infants were fatigued during the Block 3, minimising infants' task engagement. Conversely, attention attenuation during Block 2 might well be habituation after the presented stimuli has been fully encoded (e.g., Colombo & Mitchell, 2009). Taken together, it was considered most appropriate to focus our main analysis on the first 2 blocks, rather than all the 3 blocks, so the included data would better represent the neural activity whilst infants engaged in the task. Yet, at this stage, we could not fully conclude that infants' attention attenuation during Block 2 was the function of fatigue. Hence, this possibility was examined by the analysis of the association between neural dynamics between Blocks 1 and 2 of the learning phase and neural activity during the test phase in the following sections. Whilst our main analysis focused on the comparison between the Blocks 1 and 2 for the reasons described above, due to the null result from the analyses comparing Blocks 2 and 3, both the analyses comparing all the 3 blocks as well as the analyses only comparing the first 2 blocks are reported.

Table 5-5. Number of usable (i.e., artefact-free) EEG epochs

	N	M_{trials}	SD_{trials}
Block 1	32	9.11	5.99
Block 2	32	6.46	5.33
Block 3	32	5.11	5.38

Table 5-6. Attrition of EEG data

	N	$M_{attrition}$	$SD_{attrition}$
Block 1	31	0.48	0.23
Block 2	31	0.63	0.21
Block 3	27	0.70	0.21

Note. The attrition refers to the portion of the data (i.e., epochs) excluded. The sample size included in the table above only represents infants who contributed at least 1 usable trial to each block. Infants who contributed no usable trial to each block were excluded. Hence, the sample size for each block is not equal to the whole final sample of N=35.

Analysis of neural activity during the learning phase

First, we report the results of our initial analysis comparing the three blocks during the learning phase in terms of neural activity in theta-, alpha- and upper-alpha-bands. Following this, we report our main analysis focusing on the first two blocks only, as the data suggested infants' engagement to the task attenuated during the third block. After reporting these results on neural activity during the learning phase, we report the results of the analyses investigating the relationship between the learning phase and the subsequent test phases.

The normality assumption for a statistical assessment was tested using Shapiro-wilk test. Results were summarised in Tables 5-7 and 5-8. The results suggested that the data for alpha-band oscillatory activity during Block 1 and upper-alpha-band oscillatory activity during Block 2 from the sample of 20 infants did not meet the assumption of normality necessary for a parametric test. Therefore, non-parametric comparisons were used for a statistical assessment involving these datasets.

Table 5-7. Results of Shapiro-wilk test assessing the normality of the data for the initial analysis (n=15), comparing three blocks.

		n	*p*-value
Block 1	Theta (3-5Hz)	15	.87
	Alpha (6-9Hz)	15	.32
	Upper-alpha (9-10Hz)	15	.21
Block 2	Theta (3-5Hz)	15	.07
	Alpha (6-9Hz)	15	.14
	Upper-alpha (9-10Hz)	15	.09
Block 3	Theta (3-5Hz)	15	.13
	Alpha (6-9Hz)	15	.25
	Upper-alpha (9-10Hz)	15	.26

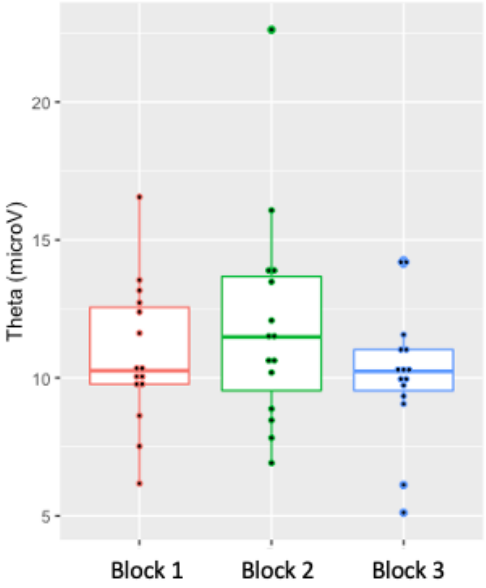
Table 5-8. Results of Shapiro-wilk test assessing the normality of the data for the analysis (N=20), comparing the first two blocks.

		n	*p*-value
Block 1	Theta (3-5Hz)	20	.99
	Alpha (6-9Hz)	20	.03
	Upper-alpha (9-10Hz)	20	.08
Block 2	Theta (3-5Hz)	20	.13
	Alpha (6-9Hz)	20	.06
	Upper-alpha (9-10Hz)	20	.02

Theta-band oscillatory activity

First, we analysed how the power of theta oscillatory activity changed across the three blocks of the learning phase, using a repeated-measures ANOVA taking Block as a within-subject measure. Theta-band oscillatory power was analysed as an index of infants' information encoding (Begus & Bonawitz, 2020). The results of the ANOVA indicated no significant difference across blocks in terms of theta-band activity ($F(2,39) = .078$, $p = .925$, $\eta^2 = 0.0035$, Cohen's $f_{\text{partial}} = 0.06$). The descriptive statistics are given in Table 5-9.

Figure 5-7. Plot showing absolute power of theta oscillatory activity across blocks.



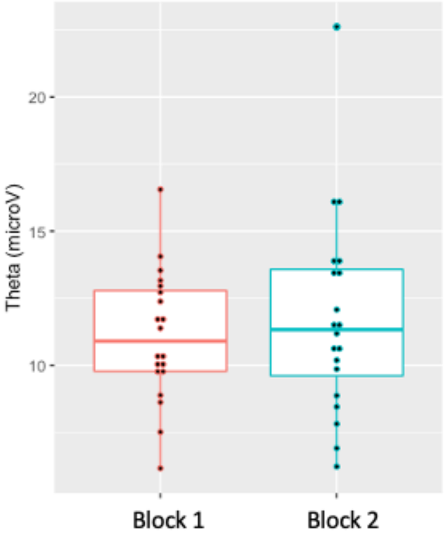
Note. The dots show each datapoint. The box and whiskers plots illustrate the maximum (the far top of the 'whisker') and minimum values (the far bottom of the 'whisker') the first (the upper boundary of the 'box') and third quartiles (the bottom boundary of the 'box'), as well as the median of the data (the line in the centre of the 'box'). As there were outliers in Blocks 2 and 3, the top and/or bottom of the 'whisker' is drawn to 1.5 x interquartile.

Whilst the repeated-measures ANOVA did not indicate a significant difference across blocks, a visual inspection of the behavioural data (video recording) suggested that there was attenuation of infants' attention during Block 3 (see previous Behavioural Analysis). Therefore, it was considered to be more appropriate to only focus on the first two blocks when examining infants' learning activity.

A paired-sample *t*-test examining the difference between Blocks 1 and 2 was performed to investigate whether there was any difference between Blocks 1 and 2 in terms of theta-band activity. For this analysis, infants who contributed to more than five usable epochs to each of the Blocks 1 and 2 were included, making the sample size *n*= 20. The descriptive statistics are summarised in Table 5-10.

A paired *t*-test exploring the difference between Blocks 1 and 2 suggested that there was no significant difference between Blocks 1 and 2 (*t*(19)= -1.01, *p* = .32,
Cohen's *d* = -0.23, 95%*CI* [-0.68, 0.23]).

Figure 5-8. Plot showing absolute power of theta oscillatory activity across blocks.



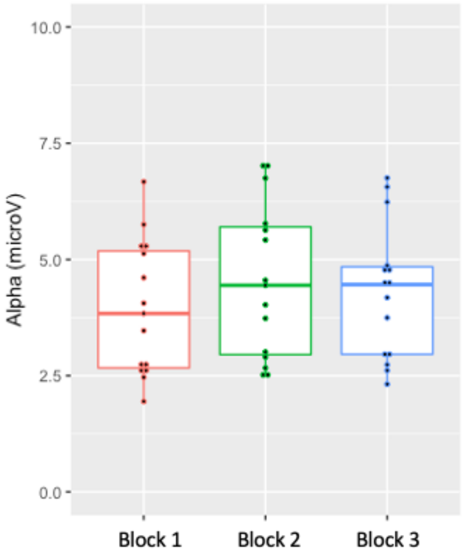
Note. The dots show each datapoint. The box and whiskers plots illustrate the maximum (the far top of the 'whisker') and minimum values (the far bottom of the 'whisker') the first (the upper boundary of the 'box') and third quartiles (the bottom boundary of the 'box'), as well as the median of the data (the line in the centre of the 'box'). As there were outliers in Block 2, the top of the "whisker" is drawn to 1.5 x interquartile.

Alpha-band oscillatory activity during the learning phase

Second, the development of power in alpha-band oscillatory activity was examined using a repeated-measures ANOVA taking Block as a within-subject measure. Alpha-band oscillatory power was assessed as an index of infants' attention (Hoehl, Michel, et al., 2014; Michel et al., 2015).

The result showed no significant difference in alpha-band oscillatory activity across blocks ($F(2,39) = .15$, $p=.86$, $\eta^2 = 0.0072$, Cohen's $f_{\text{partial}} = 0.09$). The descriptive statistics are given in Table 5-9.

Figure 5-9. Plot showing absolute power of alpha oscillatory activity across blocks.

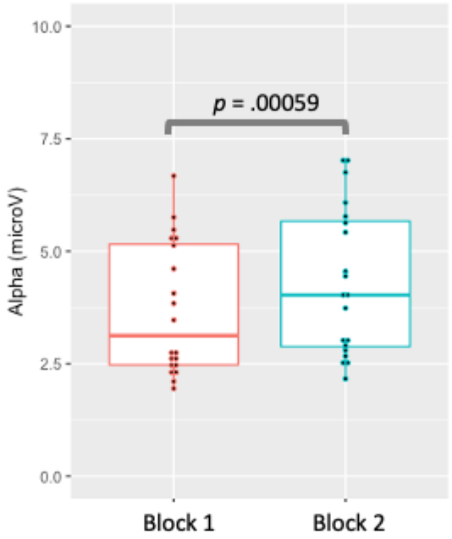


Note. The dots show each datapoint. The box and whiskers plots illustrate the maximum (the far top of the 'whisker') and minimum values (the far bottom of the 'whisker') the first (the upper boundary of the 'box') and third quartiles (the bottom boundary of the 'box'), as well as the median of the data (the line in the centre of the 'box').

In parallel with the previous analysis on theta oscillatory activity, we performed an analysis only focusing on the first two blocks. A paired-sample *t*-test examining differences between Blocks 1 and 2 in terms of alpha-band oscillatory activity was conducted. Infants who contributed at least five usable epochs to each of Blocks 1 and 2 were included as a sample, making the sample size *n*= 20. The descriptive statistics are summarised in Table 5-10.

As the data of alpha-band activity during Block 1 did not meet the assumption of normality, a non-parametric Wilcoxon signed-rank test was used to explore the difference between Blocks 1 and 2. The result suggested that alpha-band power was greater in Block 2 than Block 1 (*p* =.00059, effect size *r*= .72).

Figure 5-10. Plot showing absolute power of alpha oscillatory activity across blocks.



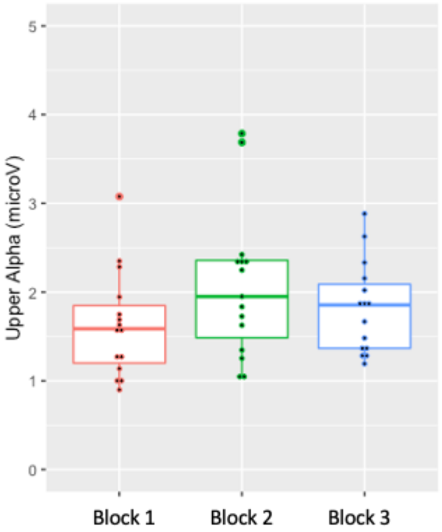
Note. The dots show each datapoint. The box and whiskers plots illustrate the maximum (the far top of the 'whisker') and minimum values (the far bottom of the 'whisker') the first (the upper boundary of the 'box') and third quartiles (the bottom boundary of the 'box'), as well as the median of the data (the line in the centre of the 'box').

Upper-alpha-band activity during the learning phase

A repeated-measures ANOVA taking Block as a within-subject measure was performed to assess whether the neural activity in the upper-alpha band differed across Blocks. Upper-alpha-band oscillatory power was used as an index of infants' learning of semantic word-object associations (Kaduk, 2016).

The result indicated no significant difference across Blocks ($F(2,39) = .524$, $p = .596$, $\eta^2 = 0.02$, Cohen's $f_{\text{partial}} = 0.16$). The descriptive statistics are given in Table 5-9.

Figure 5-11. Plot showing absolute power of upper-alpha oscillatory activity across blocks.

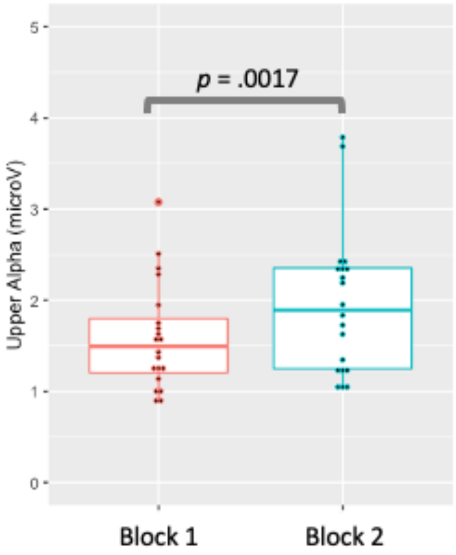


Note. The dots show each datapoint. The box and whiskers plots illustrate the maximum (the far top of the 'whisker') and minimum values (the far bottom of the 'whisker') the first (the upper boundary of the 'box') and third quartiles (the bottom boundary of the 'box'), as well as the median of the data (the line in the centre of the 'box'). As there were outliers in Blocks 1 and 2, the top and/or bottom of the 'whisker' is drawn to 1.5 x interquartile.

In parallel with the previous analyses, a paired-sample *t*-test was performed to investigate whether there was any difference between Blocks 1 and 2 in terms of the activity in the upper bound of alpha band. Infants who contributed to more than five usable epochs to each of Blocks 1 and 2 were included, making the sample size for this analysis *n*= 20. The descriptive statistics are summarised in Table 5-10.

As the data of upper-alpha-band activity during Block 1 did not meet the assumption of normality, a non-parametric Wilcoxon signed-rank test was used to explore the difference between Blocks 1 and 2. The result suggested that alpha-band power significantly increased from Blocks 1 to 2 (*p* = .0017, effect size *r*=.67).

Figure 5-12. Plot showing absolute power of upper-alpha oscillatory activity across blocks.



Note. The dots show each datapoint. The box and whiskers plots illustrate the maximum (the far top of the 'whisker') and minimum values (the far bottom of the 'whisker') the first (the upper boundary of the "box') and third quartiles (the bottom boundary of the 'box'), as well as the median of the data (the line in the centre of the 'box'). As there were outliers in Block 1, the top of the 'whisker' is drawn to 1.5 x interquartile.

Table 5-9. Descriptive statistics for the comparison across three conditions. All values are in absolute power (microV/Hz).

		Theta		**Alpha**		**Upper-alpha**	
	N	***M***	***SD***	***M***	***SD***	***M***	***SD***
Block 1	15	10.85	2.59	3.95	1.44	1.64	0.6
Block 2	15	11.91	3.89	4.53	1.65	2.07	0.83
Block 3	15	10.15	2.39	4.3	1.44	1.82	0.52

Table 5-10. Descriptive statistics for the comparison between two conditions.

		Theta		**Alpha**		**Upper-alpha**	
	N	***M***	***SD***	***M***	***SD***	***M***	***SD***
Block 1	20	11.09	2.43	3.69	1.47	1.59	0.59
Block 2	20	11.79	3.75	4.3	1.64	1.96	0.79

Neural trajectory from information encoding to knowledge consolidation

Bayesian linear mixed-effects models (LMMs) assessed how infants' neural activity during the learning phase (conducted on Day 1) predict their neural activity during the test phase immediately after the learning phase ('Day 1 test') and the test phase that happened with a 24h delay ('Day 2 test').

In the LMMs, the outcome variable, plotted on the y-axis in the following figures, was a difference in neural activity (e.g., ERP amplitude, average power in a certain frequency band) between the congruent and incongruent conditions on Day 1 and Day 2 tests (*'difference score'* hereafter). The value zero of this variable indicates that infants did not reliably discriminate between the congruent and incongruent conditions on the test phase, meaning that it is unlikely that infants reliably learned (or hold the memory of) new word-object pairs. The direction (positive or negative) of this variable is also important: as is described below, for instance, theta-band power has been found to increase when infants are presented with unexpected events as compared to expected events (Köster et al., 2019, 2021). Hence, when the difference score was taken by subtracting the power in the incongruent condition from the congruent condition (i.e., congruent minus incongruent), the negative value of this variable should indicate infants' successful learning whereas the positive value should indicate otherwise. Therefore, whether the difference score is positive or negative should be taken into account when interpreting the results according to the literature suggesting the function of each neural measure included in the analysis.

The predictor variable, plotted on the x-axis in the following figures, was a difference in neural activity (i.e., average power in a certain frequency band) between Blocks 1 and 2 during the learning phase which took place on Day 1. As we reported in the previous section, such differences could well indicate infants' cognitive processes at play during their word learning during a social interaction. More specifically, an increase in alpha-band power from Blocks 1 to 2 could be an indication of infants' attention attenuation, reflecting their habituation due to learning (Colombo & Mitchell, 2009; Nelson, 1995; Turk-Browne et al., 2006). Therefore, we hypothesised that the positive value of this predictor variable (i.e., alpha-band power), reflecting the increase in power from Blocks 1 to 2, would indicate infants' learning. Our previous analysis also found a significant increase in upper-alpha-band power from Blocks 1 to 2. Yet, there were no known previous reports identifying the function of such temporal dynamics in upper-alpha-band power during a learning event. Hence, it was challenging to form an *a priori* hypothesis. However, given an increase in alpha-band power from Blocks 1 to 2 was observed whilst upper-alpha-band power also increased from Blocks 1 to 2 and that infants' attention is closely tied to memory and learning (de Haan, 2013), we expected that an increase in upper-alpha-band power also reflects infants' learning of semantic information (i.e., the meaning of presented novel words). These mean that a positive value of this predictor variable (i.e., upper-alpha-band power) would indicate infants' learning during the learning event that took place on Day 1, whereas a negative value should indicate otherwise.

Predictor variables (theta-, alpha- and upper-alpha-band oscillatory power measured during the learning phase) did not violate the assumption of independence (*r* ranging from -0.2 to 0.23, *ps* >0.24). We fitted separate Bayesian linear mixed-effects models for each outcome variable using brms (Bürkner, 2017, 2018) and bayestestR (Makowski et al., 2019) packages in R (R Core Team, 2014). A multivariate LMM assessing all the four outcome variables were also examined but the model did not converge. This was most likely due to the small sample size of the current analysis for the complex model structure including four outcomes and four predictors. We also fitted a frequentist LMMs using lme4 function in lme4 package (Bates, Maechler, Bolker, & Walker, 2015), but the results suggested the model was a singular fit. The R script for the analysis is available on https://osf.io/p5bq8/

Theta-band power during the test phases and the neural activity during the learning phase

To assess how the neural activity during the learning phase (Day 1) was associated with theta oscillatory activity during the test phases on Days 1 and 2, we built a Bayesian LMM of the theta-band power difference between the congruent and incongruent conditions on the test phase as a function of test date, theta-, alpha- and upper-alpha-band power difference between the learning blocks 1 and 2, including interaction terms.

The main effect of 'test date' would indicate that the degree to which theta-band power differentiated between congruent and incongruent conditions of the test phases was different between Days 1 and 2 test phases. The examination of test date would act as a replication of the analyses reported by Kaduk (2016). Hence, we did not expect to observe this effect given the null result reported by Kaduk (2016). The main effect of theta-, alpha-, and upper-alpha-band power difference between Blocks 1 and 2 of the learning phase would indicate that certain dynamics of neural activity in each frequency band (e.g., an increase or decrease in theta-, alpha- and upper-alpha-band power from Blocks 1 to 2 of the learning phase taking place on Day 1 predict infants' learning (i.e., significant difference in theta-band activity between congruent and incongruent conditions of the test phase) both on Days 1 and 2 to a comparable degree. Significant interaction identified would indicate that the direction and/or the degree of the association between the neural activity difference between Blocks 1 and 2 of the learning phase which took place on Day 1 (i.e., increase or decrease from Blocks 1 to 2) and the neural activity during the Days 1 and 2 test phases differ between Days 1 and 2.

The model including interaction terms was considered theoretically better suited, as it allowed us to examine how the degree to which each neural activity (theta-, alpha- and upper-alpha-band power) influenced the theta-band power difference on the test phase varied across test dates (Days 1 and 2). Theta-band power during the test phase showed a negatively skewed distribution. Hence, the Bayesian LMM was fitted using brm function in the brms package (Bürkner, 2017, 2018) with a skewed normal distribution (family = "skew_normal"). Whilst it was considered as theoretically appropriate to include interaction terms in the model, the model without the interaction terms was also constructed to examine the effect of small sample size on the model. The model fit of the two models were compared using loo function (Vehtari et al., 2020; Vehtari, Gelman, & Gabry, 2017). It was suggested that the model including the interaction terms had a better fit than the model without the interaction terms (elpd difference -7.8, error 2.2). Thus, only the model including the interaction terms is reported.

The results were summarised in Table 5-11. The posterior distribution of coefficients from this model is shown in Figure 5-13. The visualisation of the main effects and interactions examined are provided in Figures 5-14, 5-15, 5-16, and 5-17. We found a main effect of upper-alpha-band activity, which indicates that upper-alpha-band power activity during the learning phase (Day 1) was 89% likely to have predicted infants' discrimination of the congruent and incongruent conditions of the test phases as reflected in differential theta-band EEG power across conditions (Days 1 and 2) ($\beta = -0.33$, error = 0.21, 95% credible intervals [-0.73; 0.09], 89% credible intervals [-0.65; -0.01]). This main effect of upper-alpha-band power during the learning phase (Day 1) on differences in theta-band neural activity between the congruent and incongruent conditions of the test phases (Days 1 and 2) suggests that infants who showed a more prominent upper-alpha-band power suppression during Block 1 as compared to Block 2 of the learning phase (Day 1) showed a larger theta-band power for the incongruent than the congruent condition on both Days 1 and 2 test phases.

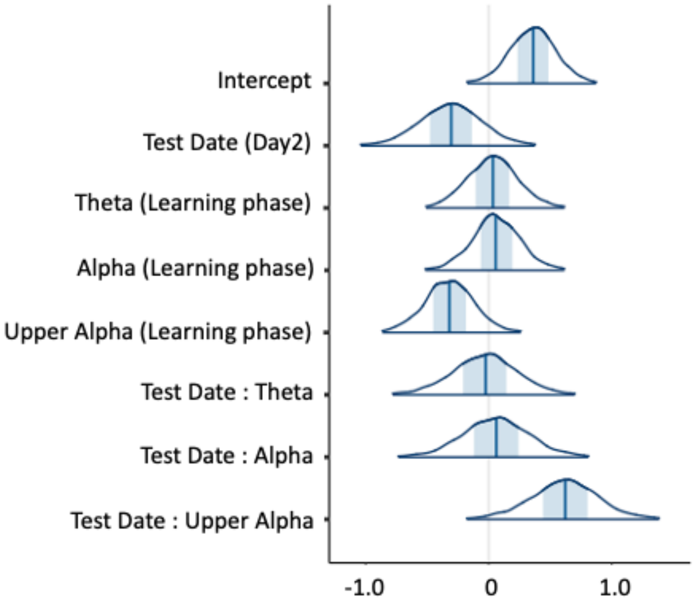
The results of this analysis also found a significant interaction between the upper-alpha-band activity during the learning phase and the test date (β = 0.63, error = 0.28, 95% credible intervals [0.09; 1.07]). The interpretation of the interaction in LMMs is dependent on the main effect (Brown, 2021). The coefficient estimate for the upper-alpha-band activity during the learning phase (Day 1) was negative (β = -0.33, error = 0.21, 95% credible intervals [-0.73, 0.09]), whereas the coefficient estimate for the interaction term was positive (β = 0.63). This indicates that, whilst the association between the upper-alpha-band activity during the learning phase (Day 1) and the theta-band activity during the test phase was negative on the Day 1 test phase, the direction became less negative (i.e., more positive) on the Day 2 test phase. This means that the degree to which upper-alpha-band power activity during the learning phase (Day 1) predicted the learning outcome (i.e., theta-band power differences between the congruent and incongruent conditions during the test phases on Days 1 and 2) became smaller on Day 2 as compared to Day 1.

Table 5-11. Bayesian linear mixed-effects model of the theta-band frequency difference score on the test phases as a function of test date, theta-, alpha- and upper-alpha-band power difference between the learning Blocks 1 and 2.

Group-Level Effects:				
Participant	(Number of levels: 14)			
	Estimate	**Est.Error**	**l-95% CI**	**u-95% CI**
sd (Intercept)	0.18	0.14	0.01	0.53

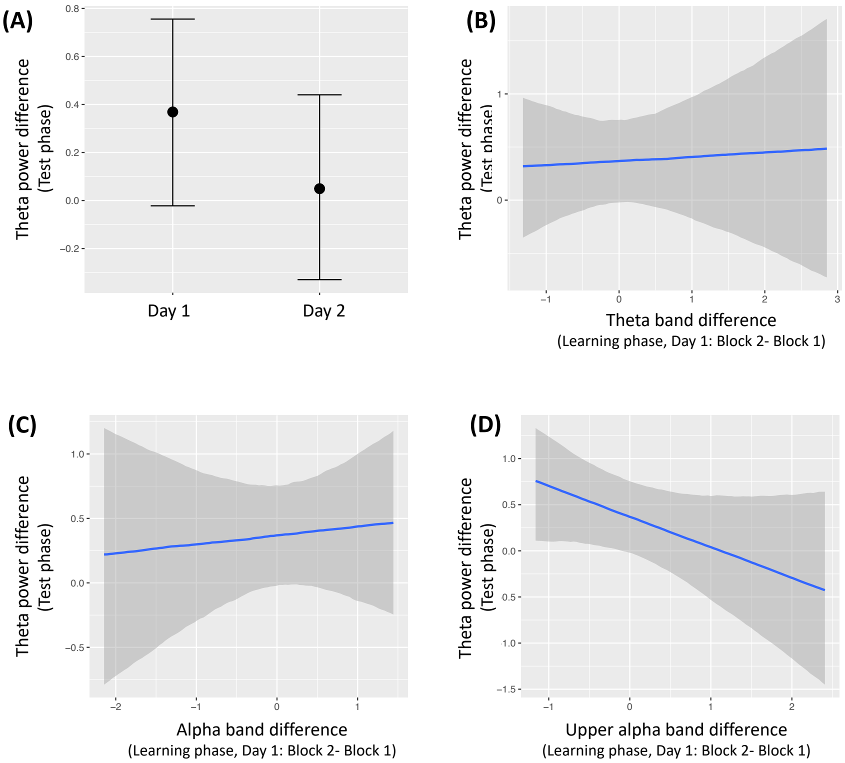
Population-Level Effects:				
	Estimate	**Est.Error**	**l-95% CI**	**u-95% CI**
Intercept	0.37	0.19	-0.02	0.76
TestDate2	-0.32	0.26	-0.84	0.21
ThetaDiff	0.04	0.21	-0.36	0.45
AlphaDiff	0.07	0.21	-0.35	0.48
UppAlphaDiff	-0.33	0.21	-0.73	0.09
TestDate2 : ThetaDiff	-0.04	0.28	-0.58	0.51
TestDate2 : AlphaDiff	0.05	0.28	-0.52	0.60
TestDate2 : UpperAlphaDiff	0.63	0.28	0.09	1.17

Figure 5-13. Posterior distributions of coefficients (plotted on x-axis) for the intercept, each predictor and interaction.



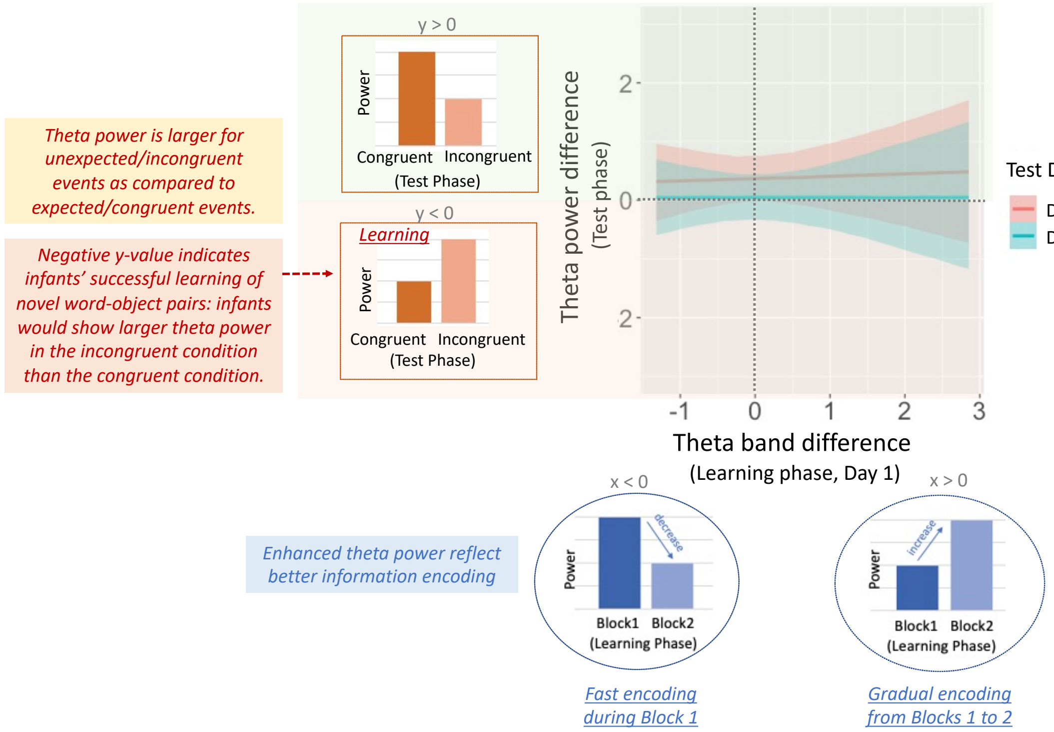
Note. The distributions represent 4000 posterior samples. The central vertical line shows the estimate mean, and the shaded region represents the 50% probability interval. The distribution tails cover the 99% probability region.

Figure 5-14. Main effect of (A) test date (with or without delay after the learning phase on Day 1), (B) theta-band power theta-band power difference between the learning blocks 1 and 2 (Day 1; M_{Block2} - M_{Block1}), (C) alpha-band power theta-band power difference between the learning blocks 1 and 2 (Day 1; M_{Block2} - M_{Block1}), and (D) upper-alpha-band power theta-band power difference between the learning blocks 1 and 2 (Day 1; M_{Block2} - M_{Block1}) on the theta-band power difference between the congruent and incongruent conditions of the test phase (y-axis: $M_{congruent}$ - $M_{incongruent}$).



Note. The "whisker" in (A) and shaded region in (B), (C), and (D) represent the 95% credible intervals. Each predictor variable (x-axis) is scaled so the mean is 0 and the standard deviation is 1. The outcome variable (y-axis) refers to a difference in theta-band power between the congruent and incongruent conditions of the test phases (Days 1 and 2). The positive value of outcome variable (y-axis) means a greater power in theta band for the congruent condition than the incongruent condition of the test phase (Day 1 or 2). Theta-band power has been reported to increase when infants are presented with unexpected events as compared to expected events (Köster et al., 2019, 2021). Hence, in the context of this study, a negative value of this outcome variable (y-axis) should indicate infants' successful detection and processing of congruent and incongruent word-object pairs. (B) The predictor (x-axis) variable refers to a difference in theta-band power between Blocks 1 and 2 of the learning phase taking place on Day 1. A positive value of predictor (x-axis) means a greater theta-band power during the learning Block 2 than Block 1. A greater theta-band power has been linked to infants' better information encoding (Begus & Bonawitz, 2020), anticipatory attention (Begus et al., 2016; Orekhova et al., 1999), as well as processing of social situations as compared to non-social situations (St. John et al., 2016). A greater theta-band power during Block 1 than 2 (i.e., decrease from Blocks 1 to 2) would indicate infants' fast encoding of the presented information in a social setting, whilst a greater power during Block 2 than 1 (i.e., increase from Blocks 1 to 2) would suggest that infants encoded information better as they heard more labelling by the experimenter. (C) The predictor (x-axis) variable refers to a difference in alpha-band power between Blocks 1 and 2 of the learning phase taking place on Day 1. A positive value represents a greater upper alpha band power during Block 2 than Block 1 (i.e., increase from Blocks 1 to 2). As suppressed alpha-band power has been linked to greater attention allocation (Hoehl et al., 2014; Klimesch, 2012; Michel et al., 2015). Therefore, an increase in alpha-band power from Blocks 1 to 2 indicates infants' attention attenuation during Block 2, suggesting infants' full encoding of the presented information during Block 1. Conversely, a decrease in alpha-band power from Blocks 1 to 2 indicates infants' sustained or increased attention during Block 2 as infants had not yet fully encoded the presented information during Block 1. (D) The predictor (x-axis) variable refers to a difference in upper-alpha-band power between Blocks 1 and 2 of the learning phase taking place on Day 1. A positive value represents a greater upper alpha band power during Block 2 than Block 1 (i.e., an increase from Blocks 1 to 2). As supressed upper-alpha-band activity has been linked to better sustention of semantic knowledge and memory (e.g., Klimsech et al., 1996), we expect that suppressed upper alpha activity should reflect better encoding of semantic information. With the literature proposing the function of neural habituation as a sign of learning (Colombo & Mitchell, 2009; Nelson, 1995; Turk-Browne et al., 2006), the increase in upper-alpha-band power from Blocks 1 to 2 (i.e., habituation from Blocks 1 to 2) could be understood as an index of infants' successful learning during Block 1.

Figure 5-15. Interaction between theta-band power difference between the learning blocks 1 and 2 (x-axis: M_{Block2} - M_{Block1}) and theta-band power difference between the congruent and incongruent conditions during the test phase (y-axis: $M_{congruent}$ - $M_{incongruent}$).


Theta power is larger for unexpected/incongruent events as compared to expected/congruent events.

Negative y-value indicates infants' successful learning of novel word-object pairs: infants would show larger theta power in the incongruent condition than the congruent condition.

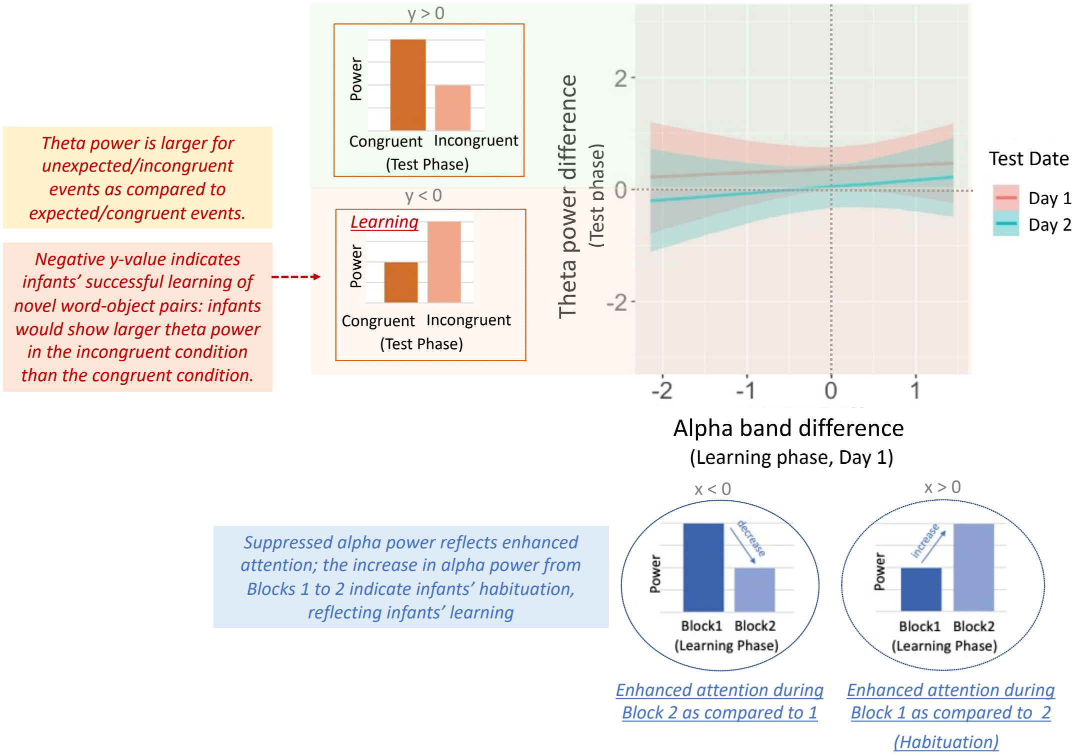
Enhanced theta power reflect better information encoding

Fast encoding during Block 1

Gradual encoding from Blocks 1 to 2

Note. The shaded region represents the 95% credible intervals. The predictor variable (x-axis) is scaled so the mean is 0 and the standard deviation is 1. The positive value of predictor (x-axis) means a greater theta-band power during the learning Block 2 than Block 1. A greater theta-band power has been linked to infants' better information encoding (Begus & Bonawitz, 2020), anticipatory attention (Begus et al., 2016; Orekhova et al., 1999), as well as processing of social situations as compared to non-social situations (St. John et al., 2016). A greater theta-band power during Block 1 than 2 (i.e., decrease from Blocks 1 to 2) would indicate infants' fast encoding of the presented information in a social setting, whilst a greater power during Block 2 than 1 (i.e., increase from Blocks 1 to 2) would suggest that infants encoded information better as they heard more labelling by the experimenter. A positive value of outcome variable (y-axis) means a greater power in theta band for the congruent than the incongruent condition of the test phase (Day 1 or 2). Theta-band power has been reported to increase when infants are presented with unexpected events as compared to expected events (Köster et al., 2019, 2021). Hence, in the context of this study, a negative value of this outcome variable (y-axis) indicates infants' successful detection and processing of congruent and incongruent word-object pairs.

Figure 5-16. Interaction between alpha-band power difference between the learning blocks 1 and 2 (x-axis: M_{Block2} - M_{Block1}) and theta-band power difference between the congruent and incongruent conditions during the test phase (y-axis: $M_{congruent}$ - $M_{incongruent}$).



Theta power is larger for unexpected/incongruent events as compared to expected/congruent events.

Negative y-value indicates infants' successful learning of novel word-object pairs: infants would show larger theta power in the incongruent condition than the congruent condition.

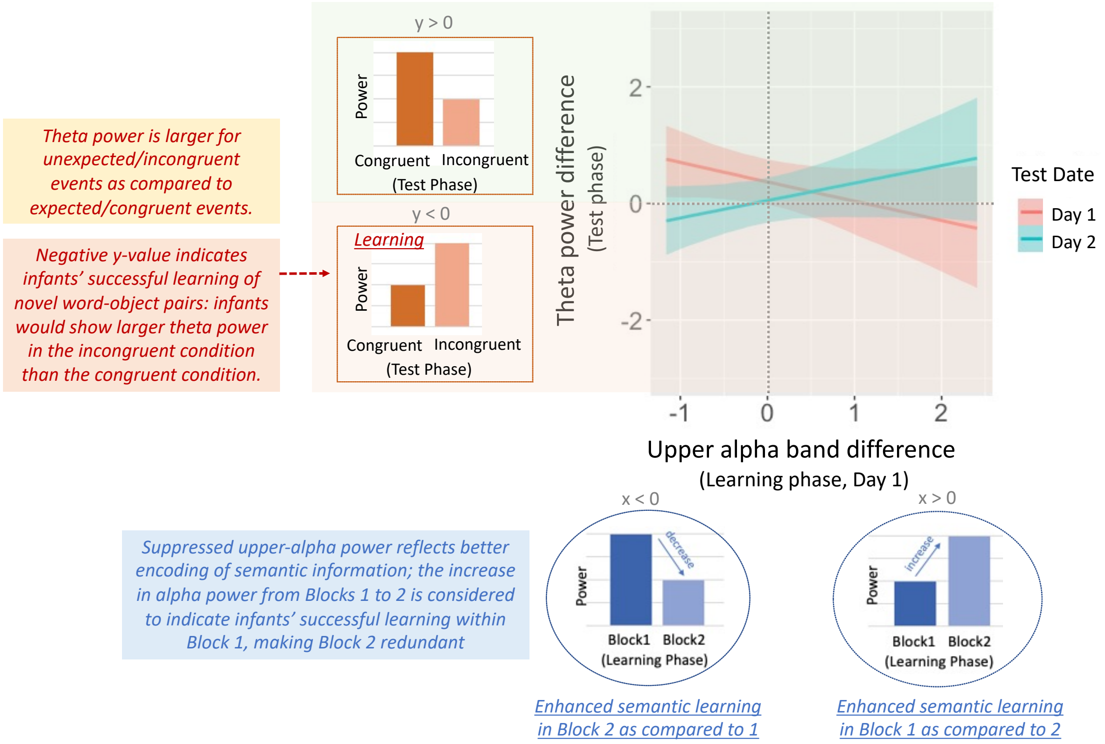
Suppressed alpha power reflects enhanced attention; the increase in alpha power from Blocks 1 to 2 indicate infants' habituation, reflecting infants' learning

Enhanced attention during Block 2 as compared to 1

Enhanced attention during Block 1 as compared to 2 (Habituation)

Note. The shaded region represents the 95% credible intervals. The predictor variable (x-axis) is scaled so the mean is 0 and the standard deviation is 1. The positive value of predictor (x-axis) means a greater alpha-band power during the learning Block 2 than Block 1. As suppressed alpha-band power has been linked to greater attention allocation (Hoehl et al., 2014; Klimesch, 2012; Michel et al., 2015). Therefore, an increase in alpha-band power from Blocks 1 to 2 indicates infants' attention attenuation during Block 2, suggesting infants' full encoding of the presented information during Block 1. Conversely, a decrease in alpha-band power from Blocks 1 to 2 indicates infants' sustained or increased attention during Block 2 as infants had not yet fully encoded the presented information during Block 1. A positive value of outcome variable (y-axis) means a greater power in theta band for the congruent condition than the incongruent condition of the test phase (Day 1 or 2). Theta-band power has been reported to increase when infants are presented with unexpected events as compared to expected events (Köster et al., 2019, 2021). Hence, in the context of this study, a negative value of this outcome variable (y-axis) should indicate infants' successful detection and processing of congruent and incongruent word-object pairs.

Figure 5-17. Interaction between upper-alpha-band power difference between the learning blocks 1 and 2 (x-axis: M_{Block2} - M_{Block1}) and theta-band power difference between the congruent and incongruent conditions during the test phase (y-axis: $M_{congruent}$ - $M_{incongruent}$).



Theta power is larger for unexpected/incongruent events as compared to expected/congruent events.

Negative y-value indicates infants' successful learning of novel word-object pairs: infants would show larger theta power in the incongruent condition than the congruent condition.

Suppressed upper-alpha power reflects better encoding of semantic information; the increase in alpha power from Blocks 1 to 2 is considered to indicate infants' successful learning within Block 1, making Block 2 redundant

Enhanced semantic learning in Block 2 as compared to 1

Enhanced semantic learning in Block 1 as compared to 2

Note. The shaded region represents the 95% credible intervals. The predictor variable (x-axis) is scaled so the mean is 0 and the standard deviation is 1. The positive value of predictor (x-axis) means a greater upper-alpha-band power during the learning Block 2 than Block 1. As supressed upper-alpha-band activity has been linked to better sustention of semantic knowledge and memory (e.g., Klimsech et al., 1996), we expect that suppressed upper alpha activity reflects better encoding of semantic information. Consistently, the literature proposing the function of neural habituation as a sign of learning (Colombo & Mitchell, 2009; Nelson, 1995; Turk-Browne et al., 2006) also suggests that an increase in upper-alpha-band power from Blocks 1 to 2 (i.e., habituation from Blocks 1 to 2) may well reflect infants' successful learning during Block 1. A positive value of outcome variable (y-axis) means a greater power in theta band for the congruent condition than the incongruent condition of the test phase (Day 1 or 2). Theta-band power has been reported to increase when infants are presented with unexpected events as compared to expected events (Köster et al., 2019, 2021). Hence, in the context of this study, a negative value of this outcome variable (y-axis) should indicate infants' successful detection and processing of congruent and incongruent word-object pairs.

Upper-alpha-band power during the test phases and the neural activity during the learning phase

To assess the relationship between the neural activity during the learning phase and the upper-alpha-band activity during the test phase, we built a Bayesian LMM of the upper-alpha-band frequency difference score on the test phases as a function of test date, theta-, alpha- and upper-alpha-band power difference between the learning blocks 1 and 2. The Bayesian LMM was fitted using brm function using a default prior and gaussian distribution (Bürkner, 2017, 2018).

The main effect of 'test date' would indicate that the degree to which upper-alpha-band power differed between congruent and incongruent conditions of the test phases was different between Days 1 and 2 test phases. Similar to the previous analysis focusing on theta-band power difference across conditions during the test phases, the examination of the test date effect was considered to be a replication of the analyses reported by Kaduk (2016), and would enable us to validate that we only observe the neural signs of infant learning (i.e., discrimination between congruent and incongruent object-label pairs) only on Day 2 but not Day 1. The main effect of theta-, alpha-, and upper-alpha-band power difference between Blocks 1 and 2 of the learning phase in this analysis would indicate that certain dynamics of oscillatory activity in each frequency band (e.g., an increase or decrease in theta-, alpha- and upper-alpha-band power from Blocks 1 to 2 of the learning phase taking place on Day 1) predicted infants' learning (i.e., (no) significant difference in neural activity in upper alpha band between congruent and incongruent conditions of the test phase) both on Days 1 and 2 to an equivalent degree. As Kaduk (2016) reported that infants showed a more suppressed (i.e., lower) upper-alpha-band oscillatory activity in the congruent condition than the incongruent condition of the test phase on Day 2 but not on Day 1, we expected to observe the main effect of test date and/or interaction between test date and other predictor variables. If the interaction was found, it would indicate that specific predictor variable(s) drawn from the neural activity during the learning phase (Day 1) predict infant learning assessed on the test phases on Days 1 and 2, but the direction and/or extent of the prediction differ between Days 1 and 2 test phases.

The model including interaction terms was theoretically considered more optimal. Yet, due to the small sample size, the model fit was assessed between models with and without interaction terms. The assessment was performed using loo package (Vehtari et al., 2020). This indicated that the difference in the fitness between the two models was negligible (elpd difference -0.3, standard error difference 1.8). Hence, the model including interaction terms is reported.

The results of this analysis are summarised in Table 5-12. The posterior distribution of coefficients from this model is shown in Figure 5-18. The visualisation of the main effects and interactions examined are provided in Figures 5-19, 5-20, 5-21, and 5-22. The analysis indicated that alpha-band power during the learning phase is 89% likely to have predicted the upper-alpha-band power during the test phases (β = -0.27; error = 0.17; 95% credible interval [-0.60; 0.05], 89% credible interval [-0.53; -0.01]). This means that infants who showed a greater power during Block 2 than Block 1, indicating infant's habituation (i.e., they had fully encoded the presented information), exhibited a greater upper-alpha-band power in the incongruent condition than the congruent condition, reflecting infants' access to semantic knowledge (Kaduk, 2016).

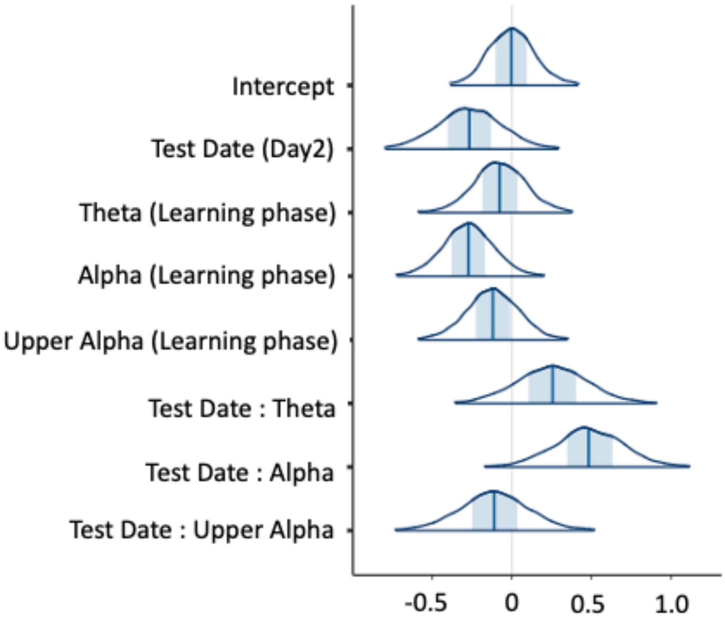
More importantly, the model found a significant interaction between the test dates and the alpha-band activity during the learning phase (β = 0.48; error = 0.23; 95% credible interval [0.03; 0.93]). As the main effect of the alpha-band power during the learning phase on the test phase suggested a negative association between the two variables (β = -0.27, error = 0.17, 95% credible intervals [-0.60, 0.05]) and the positive coefficient for this interaction of 0.48 indicates that the effect of the alpha-band power during the learning phase on the upper-alpha-band activity on the test phase was negative on the Day 1, but the effect became positive on the Day 2. A visual inspection of the graph showing this interaction (Figire 5-21) provides two indications; First, infants who showed habituation, as reflected in an increase in alpha-band power from Blocks 1 to 2 of the learning phase (Day 1), exhibited a greater upper-alpha-band power for the incongruent than the congruent condition during the Day 1 test phase, suggesting their successful learning of semantic information during the previous learning phase (Day 1) (Kaduk, 2016), whereas infants who did not show habituation, as reflected in a decrease in alpha-band power from Blocks 1 to 2 of the learning phase (Day 1), exhibited a greater upper-alpha-band power for the congruent than incongruent condition during the Day 1 test phase. Second, this pattern observed on the Day 1 test phase shifted on the Day 2 test phase; infants generally showed a greater upper-alpha-band power for the incongruent than the congruent condition during the Day 2 test phase, reflecting infants' semantic learning, regardless of the neural dynamics of the learning phase that took place on the previous day (Day 1). This suggests that infants benefited from offline consolidation that occurred between Days 1 and 2 test phases.

Table 5-12. Bayesian linear mixed-effects model of the upper-alpha-band frequency difference score on the test phases as a function of test date, theta-, alpha- and upper-alpha-band power difference between the learning Blocks 1 and 2.

Group-Level Effects:				
Participant	(Number of levels: 14)			
	Estimate	**Est.Error**	**l-95% CI**	**u-95% CI**
sd (Intercept)	0.13	0.11	0.00	0.41

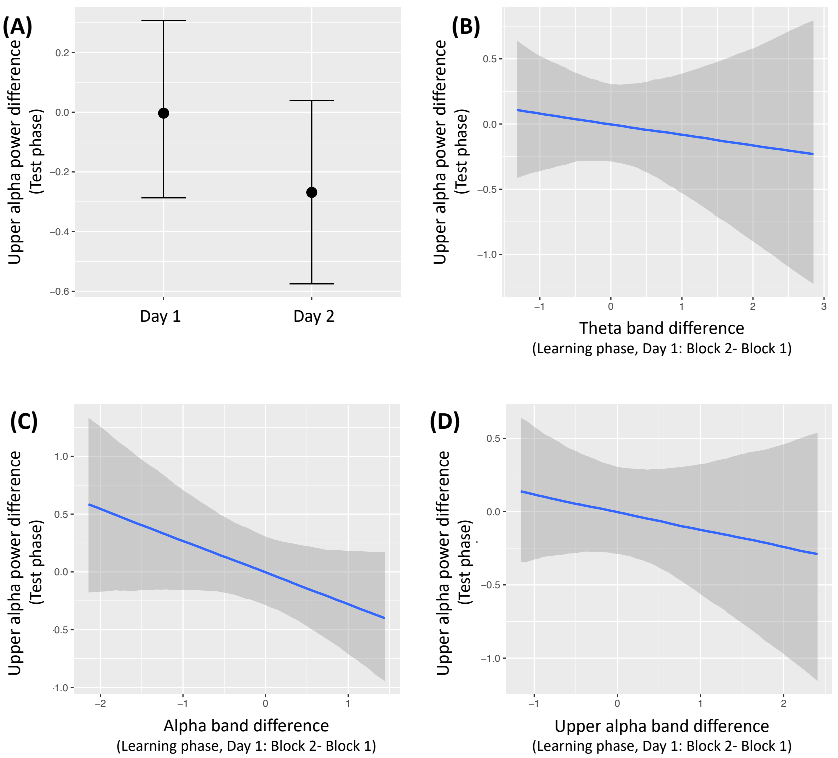
Population-Level Effects:				
	Estimate	**Est.Error**	**l-95% CI**	**u-95% CI**
Intercept	0.00	0.15	-0.29	0.31
TestDate2	-0.27	0.21	-0.68	0.14
ThetaDiff	-0.08	0.17	-0.41	0.26
AlphaDiff	-0.27	0.17	-0.60	0.05
UppAlphaDiff	-0.12	0.17	-0.46	0.22
TestDate2 : ThetaDiff	0.25	0.23	-0.19	0.71
TestDate2 : AlphaDiff	0.48	0.23	0.03	0.93
TestDate2 : UpperalphaDiff	-0.11	0.23	-0.56	0.34

Figure 5-18. Posterior distributions of intercept and coefficients for each predictor.



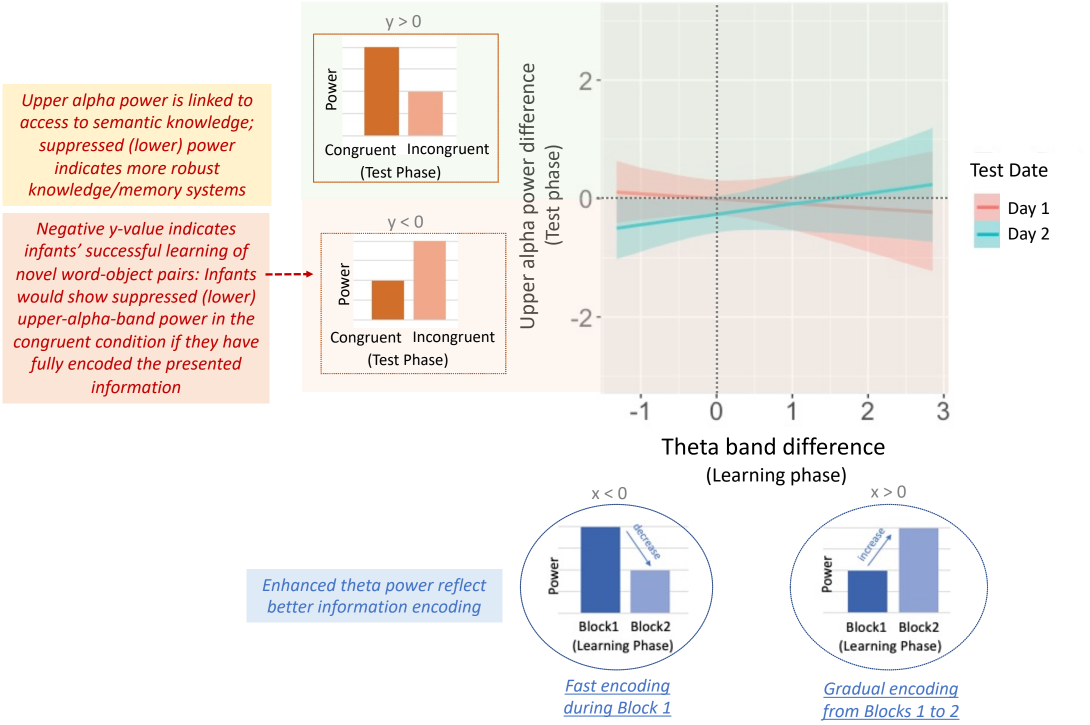
Note. The x-axis shows the estimated coefficients. The y-axis represents the density of the posterior estimates of the intercept and coefficients based on the fitted model. The distributions represent 4000 posterior samples. The central vertical line plotted in the distribution curve shows the estimate mean, and the shaded region represents the 50% probability interval. The distribution tails cover the 99% probability region.

Figure 5-19. Main effect of (A) test date (with or without delay after the learning phase on Day 1), (B) theta-band power theta-band power difference between the learning blocks 1 and 2 (Day 1; M_{Block2} - M_{Block1}), (C) alpha-band power theta-band power difference between the learning blocks 1 and 2 (Day 1; M_{Block2} - M_{Block1}), and (D) upper-alpha-band power theta-band power difference between the learning blocks 1 and 2 (Day 1; M_{Block2} - M_{Block1}) on the upper alpha-band power difference between the congruent and incongruent conditions of the test phase (y-axis: $M_{congruent}$ - $M_{incongruent}$).



Note. The "whisker" in (A) and shaded region in (B), (C), and (D) represent the 95% credible intervals. Each predictor variable (x-axis) is scaled so the mean is 0 and the standard deviation is 1. The outcome variable (y-axis) refers to a difference in upper-alpha-band power between the congruent and incongruent conditions of the test phases (Days 1 and 2). A positive value represents a greater power in upper-alpha-band for the congruent than the incongruent condition. As supressed (i.e., lower) upper-alpha-band activity has been linked to better sustention of semantic knowledge and memory (e.g., Klimsech et al., 1996), we expect that suppressed upper alpha activity reflects better encoding of semantic information. Consistently, the literature proposing the function of neural habituation as a sign of learning (Colombo & Mitchell, 2009; Nelson, 1995; Turk-Browne et al., 2006) also suggests that the increase in upper-alpha-band power from Blocks 1 to 2 (i.e., habituation from Blocks 1 to 2) could also be understood as an index of infants' successful learning during Block 1. (B) The predictor (x-axis) variable refers to a difference in theta-band power between Blocks 1 and 2 of the learning phase taking place on Day 1. A positive value of predictor (x-axis) means a greater theta-band power during the learning Block 2 than Block 1. A greater theta-band power has been linked to infants' better information encoding (Begus & Bonawitz, 2020), anticipatory attention (Begus et al., 2016; Orekhova et al., 1999), as well as processing of social situations as compared to non-social situations (St. John et al., 2016). A greater theta-band power during Block 1 than 2 (i.e., decrease from Blocks 1 to 2) would indicate infants' fast encoding of the presented information in a social setting, whilst a greater power during Block 2 than 1 (i.e., increase from Blocks 1 to 2) would suggest that infants encoded information better as they heard more labelling by the experimenter. (C) The predictor (x-axis) variable refers to a difference in alpha-band power between Blocks 1 and 2 of the learning phase taking place on Day 1. A positive value represents a greater upper alpha band power during Block 2 than Block 1 (i.e., increase from Blocks 1 to 2). As suppressed alpha-band power has been linked to greater attention allocation (Hoehl et al., 2014; Klimesch, 2012; Michel et al., 2015). Therefore, an increase in alpha-band power from Blocks 1 to 2 indicates infants' attention attenuation during Block 2, suggesting infants' full encoding of the presented information during Block 1. Conversely, a decrease in alpha-band power from Blocks 1 to 2 indicates infants' sustained or increased attention during Block 2 as infants had not yet fully encoded the presented information during Block 1. (D) The predictor (x-axis) variable refers to a difference in upper-alpha-band power between Blocks 1 and 2 of the learning phase taking place on Day 1. A positive value represents a greater upper alpha band power during Block 2 than Block 1 (i.e., an increase from Blocks 1 to 2). As supressed upper-alpha-band activity has been linked to better sustention of semantic knowledge and memory (e.g., Klimsech et al., 1996), we expect that suppressed upper alpha activity should reflect better encoding of semantic information. With the literature proposing the function of neural habituation as a sign of learning (Colombo & Mitchell, 2009; Nelson, 1995; Turk-Browne et al., 2006), the increase in upper-alpha-band power from Blocks 1 to 2 (i.e., habituation from Blocks 1 to 2) may well reflect infants' successful learning during Block 1.

Figure 5-20. Interaction between theta-band power difference between the learning blocks 1 and 2 (x-axis: $M_{Block2} - M_{Block1}$) and alpha-band power difference between the congruent and incongruent condition during the test phase (y-axis: $M_{congruent} - M_{incongruent}$).



Upper alpha power is linked to access to semantic knowledge; suppressed (lower) power indicates more robust knowledge/memory systems

Negative y-value indicates infants' successful learning of novel word-object pairs: Infants would show suppressed (lower) upper-alpha-band power in the congruent condition if they have fully encoded the presented information

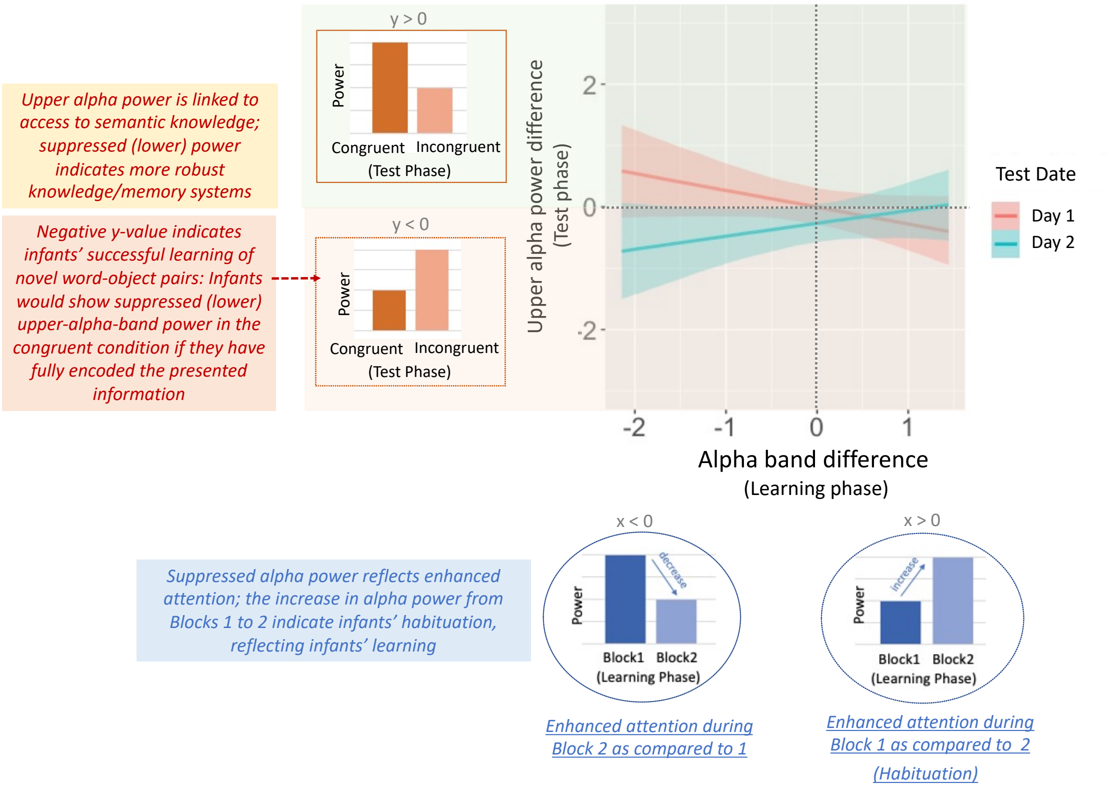
Enhanced theta power reflect better information encoding

Fast encoding during Block 1

Gradual encoding from Blocks 1 to 2

Note. *The shaded region represents the 95% credible intervals. The predictor variable (x-axis) is scaled so the mean is 0 and the standard deviation is 1. The positive value of predictor (x-axis) means a greater theta-band power during the learning Block 2 than Block 1. A greater theta-band power has been linked to infants' better information encoding (Begus & Bonawitz, 2020), anticipatory attention (Begus et al., 2016; Orekhova et al., 1999), as well as processing of social situations as compared to non-social situations (St. John et al., 2016). A greater theta-band power during Block 1 than 2 (i.e., decrease from Blocks 1 to 2) indicates infants' fast encoding of the presented information in a social setting, whilst a greater power during Block 2 than 1 (i.e., increase from Blocks 1 to 2) suggests that infants encoded information better as they heard more labelling by the experimenter. The positive value of outcome variable (y-axis) means a greater power in upper alpha band for the congruent condition than the incongruent condition. As supressed upper-alpha-band activity has been linked to better sustention of semantic knowledge and memory (e.g., Klimsech et al., 1996), we expect that suppressed upper-alpha-band activity reflects better encoding of semantic information. Consistently, the literature proposing the function of neural habituation as a sign of learning (Colombo & Mitchell, 2009; Nelson, 1995; Turk-Browne et al., 2006) also suggests that an increase in upper-alpha-band power from Blocks 1 to 2 (i.e., habituation from Blocks 1 to 2) may well reflect infants' successful learning during Block 1.*

Figure 5-21. Interaction between alpha-band power difference between the learning blocks 1 and 2 (x-axis: M_{Block2}- M_{Block1}) and alpha-band power difference between the congruent and incongruent condition during the test phase (y-axis: $M_{congruent}$ - $M_{incongruent}$).



Upper alpha power is linked to access to semantic knowledge; suppressed (lower) power indicates more robust knowledge/memory systems

Negative y-value indicates infants' successful learning of novel word-object pairs: Infants would show suppressed (lower) upper-alpha-band power in the congruent condition if they have fully encoded the presented information

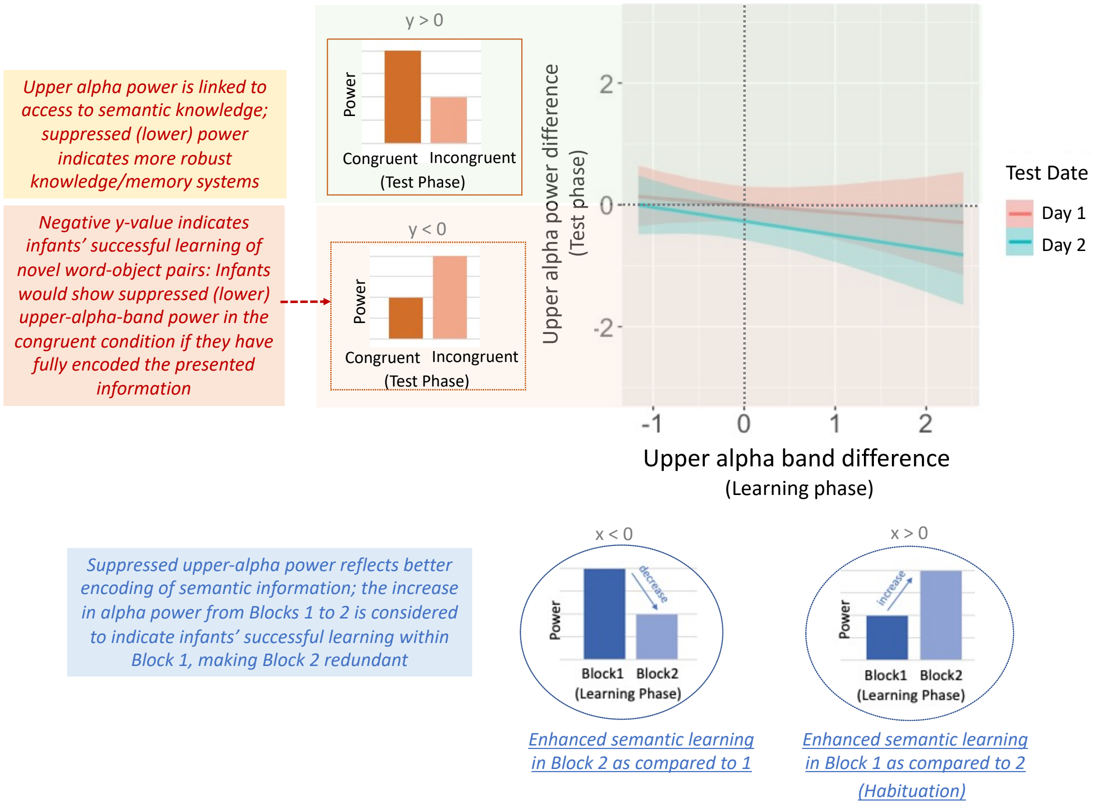
Suppressed alpha power reflects enhanced attention; the increase in alpha power from Blocks 1 to 2 indicate infants' habituation, reflecting infants' learning

Enhanced attention during Block 2 as compared to 1

Enhanced attention during Block 1 as compared to 2 (Habituation)

Note. The shaded region represents the 95% credible intervals. The predictor variable (x-axis) is scaled so the mean is 0 and the standard deviation is 1. The positive value of predictor (x-axis) means a greater alpha-band power during the learning Block 2 than Block 1. As suppressed alpha-band power has been linked to greater attention allocation (Hoehl et al., 2014; Klimesch, 2012; Michel et al., 2015). Therefore, an increase in alpha-band power from Blocks 1 to 2 indicates infants' attention attenuation during Block 2, suggesting infants' full encoding of the presented information during Block 1. Conversely, a decrease in alpha-band power from Blocks 1 to 2 indicates infants' sustained or increased attention during Block 2 as infants had not yet fully encoded the presented information during Block 1. The positive value of outcome variable (y-axis) means a greater power in upper alpha band for the congruent condition than the incongruent condition. As supressed upper-alpha-band activity has been linked to better sustention of semantic knowledge and memory (e.g., Klimsech et al., 1996), we expect that suppressed upper-alpha-band activity reflects better encoding of semantic information. Consistently, the literature proposing the function of neural habituation as a sign of learning (Colombo & Mitchell, 2009; Nelson, 1995; Turk-Browne et al., 2006) also suggests that an increase in upper-alpha-band power from Blocks 1 to 2 (i.e., habituation from Blocks 1 to 2) may well reflect infants' successful learning during Block 1.

Figure 5-22. Interaction between upper-alpha-band power difference between the learning blocks 1 and 2 (x-axis: M_{Block2}- M_{Block1}) and alpha-band power difference between the congruent and incongruent condition during the test phase (y-axis: $M_{congruent}$ - $M_{incongruent}$).



Upper alpha power is linked to access to semantic knowledge; suppressed (lower) power indicates more robust knowledge/memory systems

Negative y-value indicates infants' successful learning of novel word-object pairs: Infants would show suppressed (lower) upper-alpha-band power in the congruent condition if they have fully encoded the presented information

Suppressed upper-alpha power reflects better encoding of semantic information; the increase in alpha power from Blocks 1 to 2 is considered to indicate infants' successful learning within Block 1, making Block 2 redundant

Enhanced semantic learning in Block 2 as compared to 1

Enhanced semantic learning in Block 1 as compared to 2 (Habituation)

Note. The shaded region represents the 95% credible intervals. The predictor variable (x-axis) is scaled so the mean is 0 and the standard deviation is 1. The shaded region represents the 95% credible intervals. A positive value of predictor (x-axis) means a greater upper-alpha-band power during the learning Block 2 than Block 1(i.e., an increase from Blocks 1 to 2). As supressed upper-alpha-band activity has been linked to better sustention of semantic knowledge and memory (e.g., Klimsech et al., 1996), we expect that suppressed upper alpha activity reflects better encoding of semantic information. Consistently, the literature proposing the function of neural habituation as a sign of learning (Colombo & Mitchell, 2009; Nelson, 1995; Turk-Browne et al., 2006) also suggests that the increase in upper-alpha-band power from Blocks 1 to 2 (i.e., habituation from Blocks 1 to 2) may well reflect sign of infants' successful learning during Block 1. The positive value of outcome variable (y-axis) means a greater power in upper alpha band for the congruent condition than the incongruent condition. As supressed upper-alpha-band activity has been linked to better sustention of semantic knowledge and memory (e.g., Klimsech et al., 1996), we expect that suppressed upper-alpha-band activity reflects better encoding of semantic information. Consistently, the literature proposing the function of neural habituation as a sign of learning (Colombo & Mitchell, 2009; Nelson, 1995; Turk-Browne et al., 2006) also suggests that an increase in upper-alpha-band power from Blocks 1 to 2 (i.e., habituation from Blocks 1 to 2) may well reflect infants' successful learning during Block 1.

Nc amplitude during the test phase and the neural activity during the learning phase

To examine the relationship between the neural activity during the learning phase and the amplitude of the Nc ERP component during the test phase, we constructed a Bayesian linear LMM of the difference in the Nc mean amplitude between the congruent and incongruent conditions on the test phase as a function of test date, theta-, alpha- and upper-alpha-band power difference between the learning blocks 1 and 2.

The main effect of 'test date' would indicate that the extent to which Nc amplitude differentiated between the congruent and incongruent conditions of the test phases differed between Days 1 and 2 test phases. Similar to the previous analyses reported above, the examination of the test date effect is to replicate the analyses reported by Kaduk (2016) that we only observe the sign of infant learning reflected in neural measures (i.e., differential neural activity between the congruent and incongruent object-label pairs) only on Day 2 but not Day 1. The main effect of theta-, alpha-, and upper-alpha-band power difference between Blocks 1 and 2 of the learning phase in this analysis would indicate that dynamics of oscillatory activity in each frequency band (e.g., an increase or decrease in theta-, alpha- and upper-alpha-band power from Blocks 1 to 2 of the learning phase taking place on Day 1) predicted infants' learning (i.e., significant difference in the Nc amplitude between the congruent and incongruent conditions of the test phase) both on Days 1 and 2 to a comparable degree. As Kaduk (2016) reported that infants showed a larger-amplitude Nc in the congruent condition than the congruent condition, we expected to observe the main effect of test date and/or interaction between test date and other predictor variables. If interaction was found, it would indicate that certain neural activity during the learning phase (Day 1) predicts infant learning assessed on test phases on Days 1 and 2, but the direction and/or extent of the prediction differ between Days 1 and 2 test phases.

The model fit was assessed between models with and without interaction terms in order to evaluate the effect of the small sample size on the model. The assessment was performed using loo package (Vehtari et al., 2020, 2017). This suggested that model fitness of the two models were likely to be comparable (elpd difference -2.2, standard error difference 2.1). Since the model should consider interaction terms according to our theoretical background, only the model with interaction terms is reported.

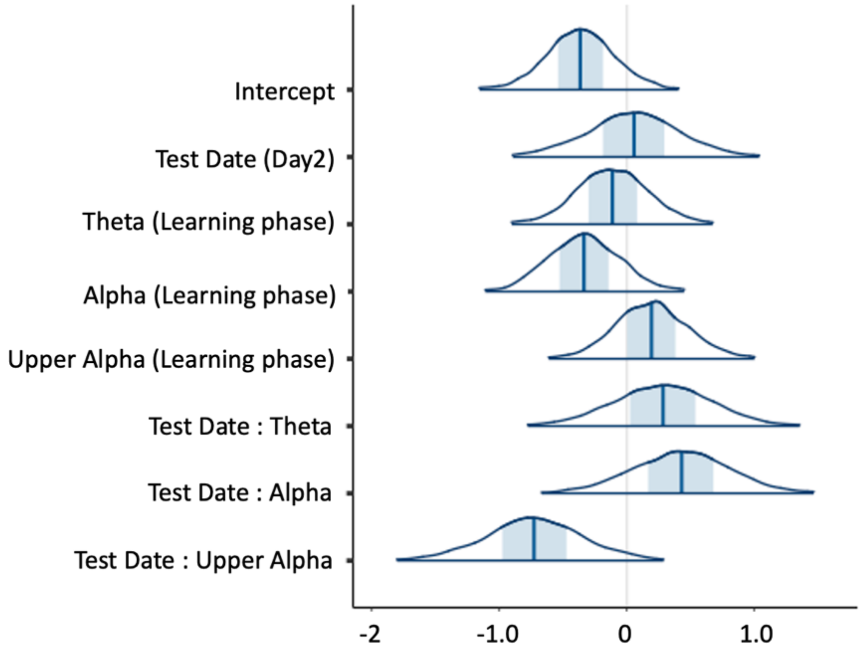
The results of this analysis are summarised in Table 5-13. The posterior distribution of coefficients from this model is shown in Figure 5-23. The visualisation of the main effects and interactions examined are provided in Figures 5-24, 5-25, 5-26, and 5-27. Despite no significant main effects, the results indicated a potential interaction between the upper-alpha-band activity during the learning phase and the test date (*β* = -0.73, error = 0.39, 95% credible intervals [-1.52; 0.02], 94% credible intervals [-1.48; -0.02]). The coefficient estimate for the upper-alpha-band activity during the learning phase was positive (*β* = 0.20, error = 0.30, 95% credible intervals [-0.37, 0.79]), whereas the coefficient estimate for the interaction term was negative (*β* = -0.73). This suggests that, whilst the association between the upper-alpha-band activity during the learning phase that took place on the Day 1 and the Nc mean amplitude on the Day 1 test phase was positive, the direction became more negative on the Day 2 test phase. A further examination of the figure (Figure 5-25) indicated that infants who exhibited a greater power in the upper-alpha frequency band during Block 2 than Block 1 of the learning phase (Day 1), reflecting their habituation in Block 2 due to learning (i.e., full information encoding) having occurred during Block 1, showed a larger amplitude of Nc for the incongruent condition than the congruent condition on the test phase during the Day 1 test phase. Conversely, infants who showed a greater upper-alpha-band power during Block 1 than Block 2 of the learning phase (Day 1), indicating continued encoding from Blocks 1 to 2, exhibited no significant difference (i.e., (close to) zero difference) in Nc amplitude on the Day 1 test phase, suggesting that they did not reliably discriminate the congruent and incongruent word-object pairs immediately after the learning phase. Critically though, this pattern changed on the Day 2 test phase: infants generally reliably showed a larger-amplitude Nc during the incongruent condition than the congruent condition, suggesting their discrimination between congruent and incongruent word-object pairs. This indicates the effect of offline consolidation that happened during the Day 1 and 2 test phases.

Table 5-13. Bayesian linear mixed-effects model of the Nc mean amplitude difference score on the test phases as a function of test date, theta-, alpha- and upper-alpha-band power difference between the learning Blocks 1 and 2.

Group-Level Effects:				
Participant	(Number of levels: 14)			
	Estimate	**Est.Error**	**l-95% CI**	**u-95% CI**
sd (Intercept)	0.24	0.20	0.01	0.73

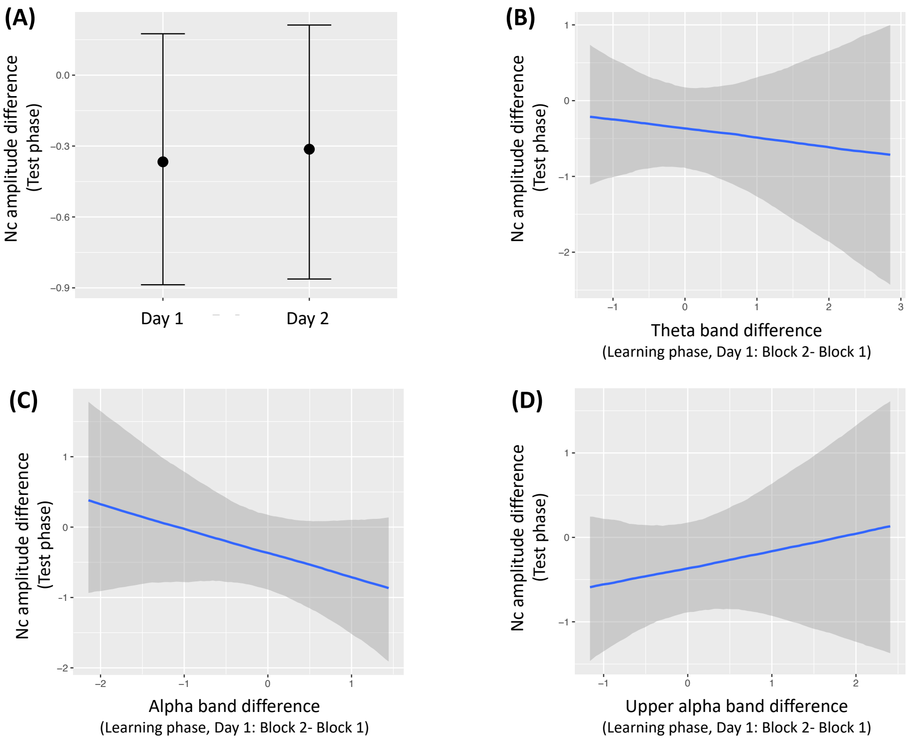
Population-Level Effects:				
	Estimate	**Est.Error**	**l-95% CI**	**u-95% CI**
Intercept	-0.36	0.28	-0.91	0.19
TestDate2	0.05	0.37	-0.70	0.78
ThetaDiff	-0.11	0.29	-0.69	0.48
AlphaDiff	-0.33	0.29	-0.87	0.25
UppAlphaDiff	0.20	0.30	-0.37	0.79
TestDate2 : ThetaDiff	0.28	0.40	-0.51	1.04
TestDate2 : AlphaDiff	0.43	0.39	-0.36	1.20
TestDate2 : UpperalphaDiff	-0.73	0.39	-1.52	0.02

Figure 5-23. Posterior distributions of intercept and coefficients for each predictor.



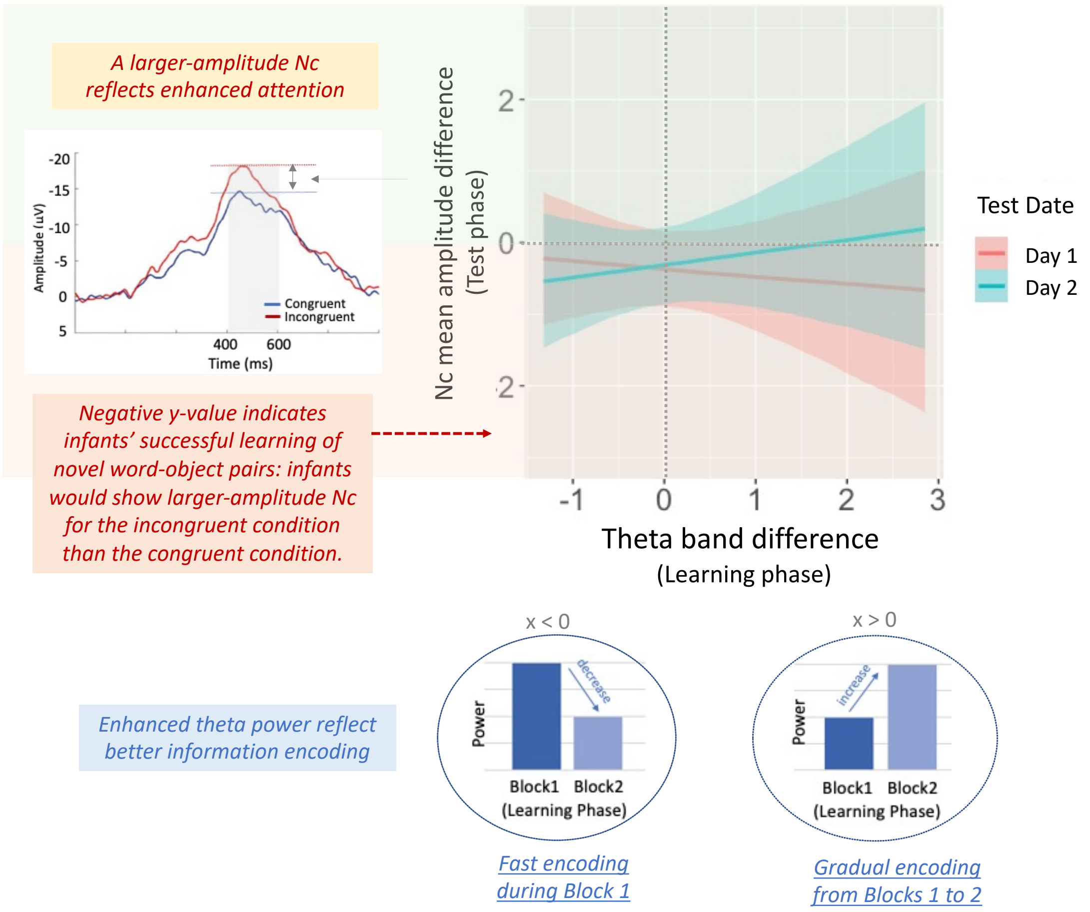
Note. The x-axis shows the estimated coefficients. The y-axis represents the density of the posterior estimates of the intercept and coefficients based on the fitted model. The distributions represent 4000 posterior samples. The central vertical line plotted in the distribution curve shows the estimate mean, and the shaded region represents the 50% probability interval. The distribution tails cover the 99% probability region.

Figure 5-24. Main effect of (A) test date (with or without delay after the learning phase on Day 1), (B) theta-band power theta-band power difference between the learning blocks 1 and 2 (Day 1; M_{Block2} - M_{Block1}), (C) alpha-band power theta-band power difference between the learning blocks 1 and 2 (Day 1; M_{Block2} - M_{Block1}), and (D) upper-alpha-band power theta-band power difference between the learning blocks 1 and 2 (Day 1; M_{Block2} - M_{Block1}) on the Nc amplitude difference between the congruent and incongruent conditions during the test phase (y-axis: $M_{incongruent}$ - $M_{congruent}$).



Note. The "whisker" in (A) and shaded region in (B), (C), and (D) represent the 95% credible intervals. Each predictor variable (x-axis) is scaled so the mean is 0 and the standard deviation is 1. The outcome variable (y-axis) refers to a difference in Nc amplitude between the congruent and incongruent conditions of the test phases (Days 1 and 2). The positive value means that Nc amplitude was larger for the congruent condition when contrasted with the incongruent condition. We expect that infants who have encoded information (i.e., learned the presented information) show enhanced attention (reflected in a larger amplitude Nc) towards the incongruent condition when contrasted to the congruent condition. Therefore, a negative value of the predictor variable (i.e., Nc amplitude difference) reflects infants' successful learning and knowledge consolidation. (B) The predictor (x-axis) variable refers to a difference in theta-band power between Blocks 1 and 2 of the learning phase taking place on Day 1. A positive value of predictor (x-axis) means a greater theta-band power during the learning Block 2 than Block 1. A greater theta-band power has been linked to infants' better information encoding (Begus & Bonawitz, 2020), anticipatory attention (Begus et al., 2016; Orekhova et al., 1999), as well as processing of social situations as compared to non-social situations (St. John et al., 2016). A greater theta-band power during Block 1 than 2 (i.e., decrease from Blocks 1 to 2) would indicate infants' fast encoding of the presented information in a social setting, whilst a greater power during Block 2 than 1 (i.e., increase from Blocks 1 to 2) would suggest that infants encoded information better as they heard more labelling by the experimenter. (C) The predictor (x-axis) variable refers to a difference in alpha-band power between Blocks 1 and 2 of the learning phase taking place on Day 1. A positive value represents a greater upper alpha band power during Block 2 than Block 1 (i.e., increase from Blocks 1 to 2). As suppressed alpha-band power has been linked to greater attention allocation (Hoehl et al., 2014; Klimesch, 2012; Michel et al., 2015). Therefore, an increase in alpha-band power from Blocks 1 to 2 indicates infants' attention attenuation during Block 2, suggesting infants' full encoding of the presented information during Block 1. Conversely, a decrease in alpha-band power from Blocks 1 to 2 indicates infants' sustained or increased attention during Block 2 as infants had not yet fully encoded the presented information during Block 1. (D) The predictor (x-axis) variable refers to a difference in upper-alpha-band power between Blocks 1 and 2 of the learning phase taking place on Day 1. A positive value represents a greater upper alpha band power during Block 2 than Block 1 (i.e., an increase from Blocks 1 to 2). As supressed upper-alpha-band activity has been linked to better sustention of semantic knowledge and memory (e.g., Klimsech et al., 1996), we expect that suppressed upper alpha activity reflects better encoding of semantic information. With the literature proposing the function of neural habituation as a sign of learning (Colombo & Mitchell, 2009; Nelson, 1995; Turk-Browne et al., 2006), the increase in upper-alpha-band power from Blocks 1 to 2 (i.e., habituation from Blocks 1 to 2) may well reflect infants' successful learning during Block 1.

Figure 5-25. Interaction between theta-band power difference between the learning blocks 1 and 2 (x-axis: M_{Block2} - M_{Block1}) and the difference in the Nc mean amplitude between the congruent and incongruent condition during the test phase (y-axis: $M_{incongruent}$ - $M_{congruent}$).



A larger-amplitude Nc reflects enhanced attention

Negative y-value indicates infants' successful learning of novel word-object pairs: infants would show larger-amplitude Nc for the incongruent condition than the congruent condition.

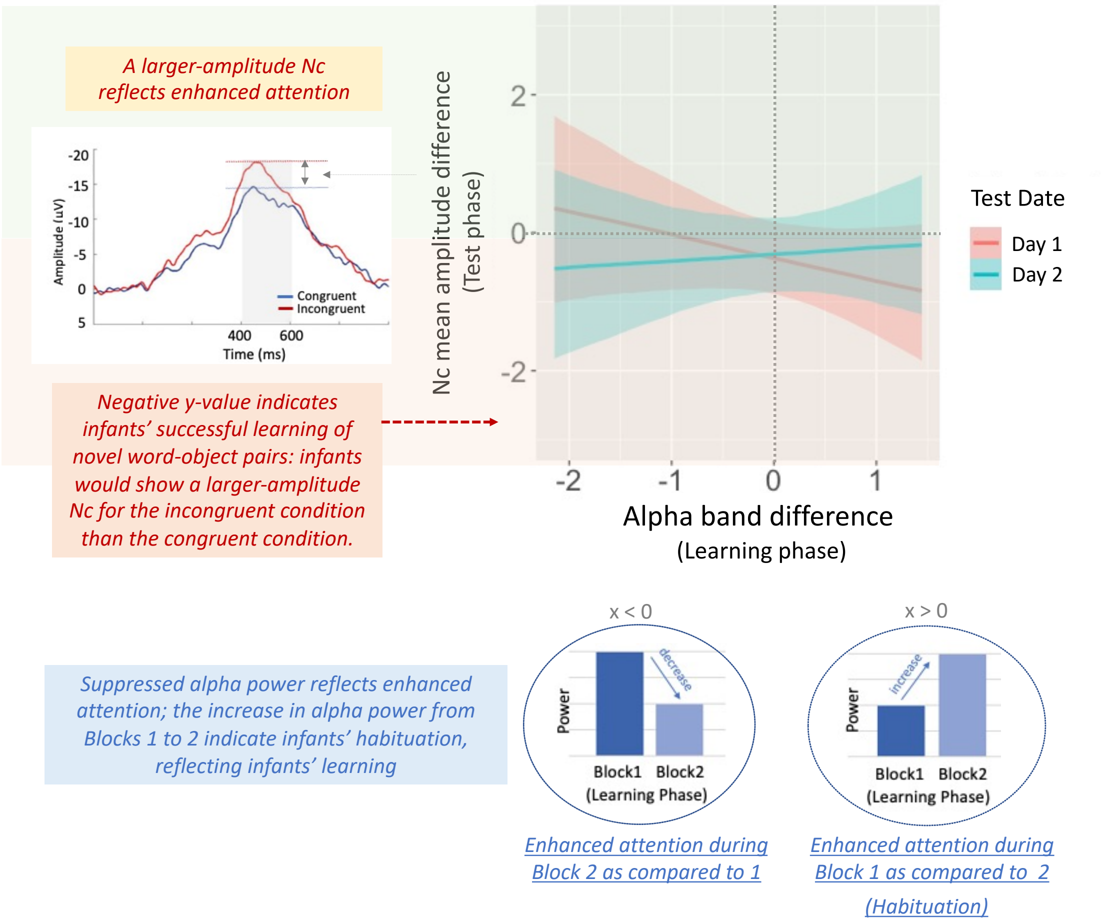
Enhanced theta power reflect better information encoding

Fast encoding during Block 1

Gradual encoding from Blocks 1 to 2

Note. The shaded region represents the 95% credible intervals. The predictor variable (x-axis) is scaled so the mean is 0 and the standard deviation is 1. The positive value of predictor (x-axis) means a greater theta-band power during the learning Block 2 than Block 1. A greater theta-band power has been linked to infants' better information encoding (Begus & Bonawitz, 2020), anticipatory attention (Begus et al., 2016; Orekhova et al., 1999), as well as processing of social situations as compared to non-social situations (St. John et al., 2016). A greater theta-band power during Block 1 than 2 (i.e., decrease from Blocks 1 to 2) indicates infants' fast encoding of the presented information in a social setting, whilst a greater power during Block 2 than 1 (i.e., increase from Blocks 1 to 2) suggests that infants encoded information better as they heard more labelling by the experimenter. The positive value of outcome variable (y-axis) means that Nc amplitude was larger for the congruent condition when contrasted with the incongruent condition. We expect that infants who have encoded information (i.e., learned the presented information) show enhanced attention (reflected in a larger amplitude Nc) towards the incongruent condition when contrasted to the congruent condition. Therefore, a negative value of the predictor variable (i.e., Nc amplitude difference) reflects infants' successful learning and knowledge consolidation.

Figure 5-26. Interaction between alpha-band power difference between the learning blocks 1 and 2 (x-axis: M_{Block2} - M_{Block1}) and the difference in the Nc mean amplitude between the congruent and incongruent condition during the test phase (y-axis: $M_{incongruent}$ - $M_{congruent}$).



A larger-amplitude Nc reflects enhanced attention

Negative y-value indicates infants' successful learning of novel word-object pairs: infants would show a larger-amplitude Nc for the incongruent condition than the congruent condition.

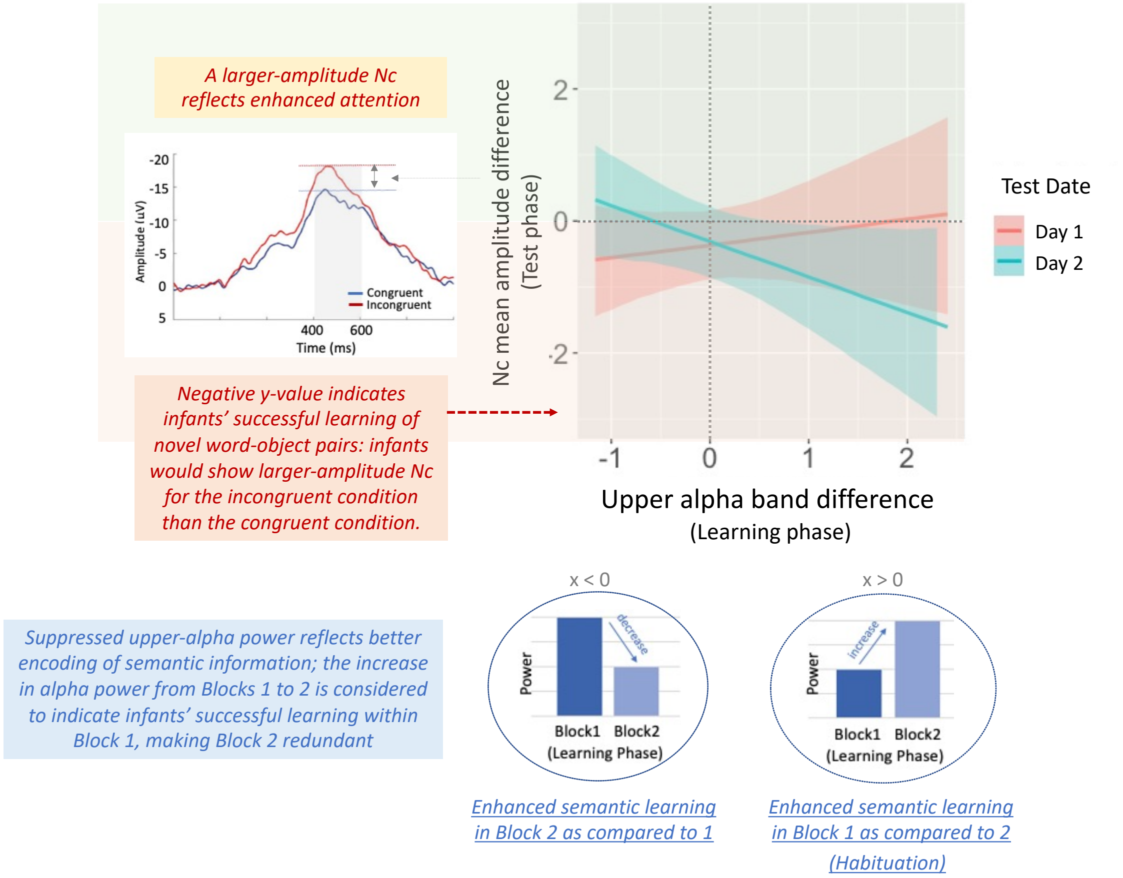
Suppressed alpha power reflects enhanced attention; the increase in alpha power from Blocks 1 to 2 indicate infants' habituation, reflecting infants' learning

Enhanced attention during Block 2 as compared to 1

Enhanced attention during Block 1 as compared to 2 (Habituation)

Note. The shaded region represents the 95% credible intervals. The predictor variable (x-axis) is scaled so the mean is 0 and the standard deviation is 1. The positive value of predictor (x-axis) means a greater alpha-band power during the learning Block 2 than Block 1. As suppressed alpha-band power has been linked to greater attention allocation (Hoehl et al., 2014; Klimesch, 2012; Michel et al., 2015). Therefore, the increase in alpha-band power from Blocks 1 to 2 indicates infants' attention attenuation during Block 2, suggesting infants' full encoding of the presented information during Block 1. Conversely, the decrease in alpha-band power from Blocks 1 to 2 indicates infants' sustained or increased attention during Block 2 as infants had not yet fully encoded the presented information during Block 1. The positive value of outcome variable (y-axis) means that Nc amplitude was larger for the congruent condition when contrasted with the incongruent condition. We expect that infants who have encoded information (i.e., learned the presented information) show enhanced attention (reflected in a larger amplitude Nc) towards the incongruent condition when contrasted to the congruent condition. Therefore, a negative value of the predictor variable (i.e., Nc amplitude difference) reflects infants' successful learning and knowledge consolidation.

Figure 5-27. Interaction between upper-alpha-band power difference between the learning blocks 1 and 2 (x-axis: M_{Block2} - M_{Block1}) and the difference in the Nc mean amplitude between the congruent and incongruent condition during the test phase (y-axis: $M_{incongruent}$ - $M_{congruent}$).



A larger-amplitude Nc reflects enhanced attention

Negative y-value indicates infants' successful learning of novel word-object pairs: infants would show larger-amplitude Nc for the incongruent condition than the congruent condition.

Suppressed upper-alpha power reflects better encoding of semantic information; the increase in alpha power from Blocks 1 to 2 is considered to indicate infants' successful learning within Block 1, making Block 2 redundant

Enhanced semantic learning in Block 2 as compared to 1

Enhanced semantic learning in Block 1 as compared to 2 (Habituation)

Note. The shaded region represents the 95% credible intervals. The predictor variable (x-axis) is scaled so the mean is 0 and the standard deviation is 1. The positive value of predictor (x-axis) means a greater upper-alpha-band power during the learning Block 2 than Block 1. As supressed upper-alpha-band activity has been linked to better sustention of semantic knowledge and memory (e.g., Klimsech et al., 1996), we expect that suppressed upper alpha activity should reflect better encoding of semantic information. The literature proposing the function of neural habituation as a sign of learning (Colombo & Mitchell, 2009; Nelson, 1995; Turk-Browne et al., 2006) also suggests that an increase in upper-alpha-band power from Blocks 1 to 2 (i.e., habituation from Blocks 1 to 2) may well reflect infants' successful learning during Block 1. The positive value of outcome variable (y-axis) means that Nc amplitude was larger for the congruent condition when contrasted with the incongruent condition. We expect that infants who have encoded information (i.e., learned the presented information) show enhanced attention (reflected in a larger amplitude Nc) towards the incongruent condition when contrasted to the congruent condition. Therefore, a negative value of the predictor variable (i.e., Nc amplitude difference) reflects infants' successful learning and knowledge consolidation.

N400 during the test phase and the neural activity during the learning phase

We constructed a Bayesian LMM of the difference in the N400 (quantified as an area under the curve on the waveform) between the congruent and incongruent conditions on the test phase, as a function of test date, theta-, alpha- and upper-alpha-band power difference between the learning blocks 1 and 2.

The main effect of 'test date' would indicate that how the N400 was observed in the test phases (i.e., the N400 should be more prominent in the incongruent condition than the congruent condition if infants learned novel words and detected semantic incongruency, based on previous literature (e.g., Reid et al., 2009)) differed between Days 1 and 2 test phases. As was the case with the previous analyses reported above, the examination of the test date effect is a replication of the analyses reported by Kaduk (2016) which suggested that the neural signs of infant learning (i.e., more prominent N400 for the incongruent condition than the congruent condition) should be only evident on Day 2 but not Day 1. The main effect of theta-, alpha-, and upper-alpha-band power difference between Blocks 1 and 2 of the learning phase in this analysis would indicate that certain dynamics of oscillatory activity in each frequency band (e.g., an increase or decrease in theta-, alpha- and upper-alpha-band power from Blocks 1 to 2 of the learning phase taking place on Day 1) predicted infants' learning (i.e., more evident N400 effect in the incongruent condition than the congruent condition) both on Days 1 and 2 to a comparable degree. As Kaduk (2016) reported a significant N400 effect which indicates infants' successful discrimination between the congruent and incongruent conditions, we expected to observe the main effect of test date and/or interaction between test date and other predictor variables. Any significant interaction identified in this analysis would indicate that the association between a predictor variable from the neural activity during the learning phase (Day 1) and the neural activity difference between the congruent and incongruent conditions of the test phases would differ between Days 1 and 2.

The model fit was assessed between models with and without interaction terms due to a potential issue with the small sample size for the complex model structure including four predictors. The model fit comparison was performed using loo package (Vehtari et al., 2020, 2017), which indicated the model without interaction terms had a better fit (elpd difference -146.9, standard error difference 5.4). Whilst the model including the interaction terms would theoretically be better justified, due to the poor fit, it is considered best to examine the model without interaction terms.

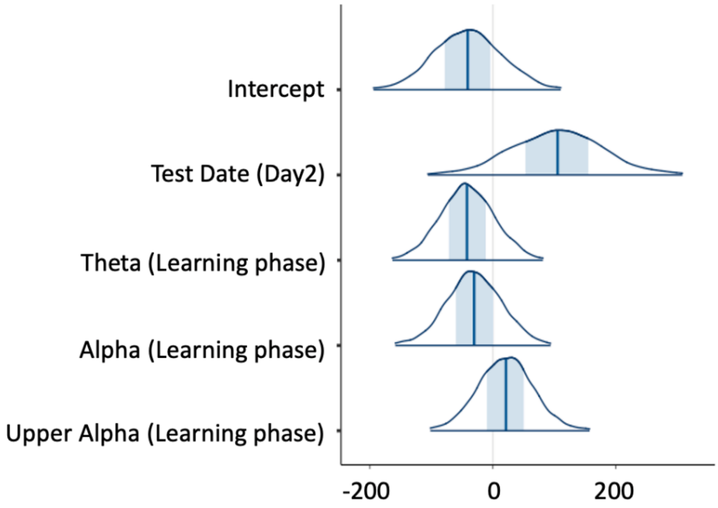
The results of this analysis are summarised in Table 5-14. The posterior distribution of coefficients from this model is shown in Figure 5-28. The main effects are visualised in Figure 5-29. The model indicated that none of the predictor variables was reliably associated with the N400 difference score during the test phases on Days 1 and 2. The strongest effect was the test date (86% credible intervals [2.39, 220.60]), which suggests that in 86% of the case, the N400 effect (i.e., more prominent N400 for the incongruent condition than the congruent condition) was more significantly evident on Day 2 than on Day1. As the N400 has been reported to be evident in response to the detection of semantic incongruency (e.g., Friedrich & Friedrici, 2008), this indicates that infants only reliably discriminated the semantic (in)congruency in the presented stimuli on Day 2, not Day 1, supporting the complementary learning account emphasising the importance of offline consolidation (Davis & Gaskell, 2009; McClelland et al., 1995; Norman & O'Reilly, 2003; O'Reilly & Norman, 2002; Weighall et al., 2017).

Table 5-14. Bayesian linear mixed-effects model of the N400 divergence score difference on the test phases as a function of test date, theta-, alpha- and upper-alpha-band power difference between the learning Blocks 1 and 2.

Group-Level Effects:				
Participant	(Number of levels: 14)			
	Estimate	**Est.Error**	**l-95% CI**	**u-95% CI**
sd(Intercept)	61.81	45.53	2.44	172.32

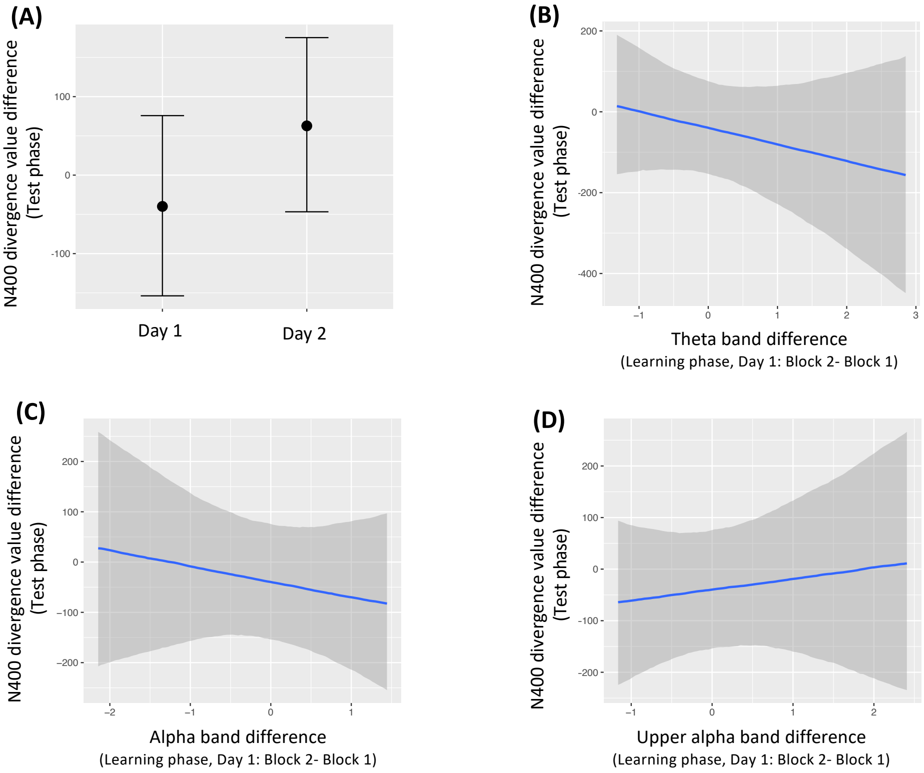
Population-Level Effects:				
	Estimate	**Est.Error**	**l-95% CI**	**u-95% CI**
Intercept	-40.55	56.53	-149.75	70.5
TestDate2	104.46	76.73	-46.04	259.17
ThetaDiff	-41.47	45.63	-130.61	46.69
AlphaDiff	-30.17	46.8	-122.78	61.96
UppAlphaDiff	21.1	46.44	-68.42	114.23

Figure 5-28. Posterior distributions of intercept and coefficients for each predictor.



Note. The x-axis shows the estimated coefficients. The y-axis represents the density of the posterior estimates of the intercept and coefficients based on the fitted model. The distributions represent 4000 posterior samples. The central vertical line plotted in the distribution curve shows the estimate mean, and the shaded region represents the 50% probability interval. The distribution tails cover the 99% probability region.

Figure 5-29. Main effect of (A) test date (with or without delay after the learning phase on Day 1), (B) theta-band power theta-band power difference between the learning blocks 1 and 2 (Day 1; M_{Block2}- M_{Block1}), (C) alpha-band power theta-band power difference between the learning blocks 1 and 2 (Day 1; M_{Block2}- M_{Block1}), and (D) upper-alpha-band power theta-band power difference between the learning blocks 1 and 2 (Day 1; M_{Block2} - M_{Block1}) on the N400 divergence value difference between the congruent and incongruent conditions during the test phase (y-axis: $M_{incongruent}$ - $M_{congruent}$).



Note. The "whisker" in (A) and shaded region in (B), (C), and (D) represent the 95% credible intervals. Each predictor variable (x-axis) is scaled so the mean is 0 and the standard deviation is 1. The outcome variable (y-axis) refers to a difference in N400 divergence value between the congruent and incongruent conditions of the test phases (Days 1 and 2) (for score calculation, see previous section or Kaduk, 2016). A positive value represents a more prominent N400 during the incongruent condition as compared to the congruent condition. N400 component has been reported upon detection and processing of semantic incongruency (e.g., Friedrich & Friedrici, 2008). Thus, a positive value indicates infants' successful semantic learning. (B) The predictor (x-axis) variable refers to a difference in theta-band power between Blocks 1 and 2 of the learning phase taking place on Day 1. A positive value of predictor (x-axis) means a greater theta-band power during the learning Block 2 than Block 1. A greater theta-band power has been linked to infants' better information encoding (Begus & Bonawitz, 2020), anticipatory attention (Begus et al., 2016; Orekhova et al., 1999), as well as processing of social situations as compared to non-social situations (St. John et al., 2016). A greater theta-band power during Block 1 than 2 (i.e., decrease from Blocks 1 to 2) would indicate infants' fast encoding of the presented information in a social setting, whilst a greater power during Block 2 than 1 (i.e., increase from Blocks 1 to 2) would suggest that infants encoded information better as they heard more labelling by the experimenter. (C) The predictor (x-axis) variable refers to a difference in alpha-band power between Blocks 1 and 2 of the learning phase taking place on Day 1. A positive value represents a greater upper alpha band power during Block 2 than Block 1 (i.e., increase from Blocks 1 to 2). As suppressed alpha-band power has been linked to greater attention allocation (Hoehl et al., 2014; Klimesch, 2012; Michel et al., 2015). Therefore, an increase in alpha-band power from Blocks 1 to 2 indicates infants' attention attenuation during Block 2, suggesting infants' full encoding of the presented information during Block 1. Conversely, a decrease in alpha-band power from Blocks 1 to 2 indicates infants' sustained or increased attention during Block 2 as infants had not yet fully encoded the presented information during Block 1. (D) The predictor (x-axis) variable refers to a difference in upper-alpha-band power between Blocks 1 and 2 of the learning phase taking place on Day 1. A positive value represents a greater upper alpha band power during Block 2 than Block 1 (i.e., an increase from Blocks 1 to 2). As supressed upper-alpha-band activity has been linked to better sustention of semantic knowledge and memory (e.g., Klimsech et al., 1996), we expect that suppressed upper alpha activity should reflect better encoding of semantic information. With the literature proposing the function of neural habituation as a sign of learning (Colombo & Mitchell, 2009; Nelson, 1995; Turk-Browne et al., 2006), the increase in upper-alpha-band power from Blocks 1 to 2 (i.e., habituation from Blocks 1 to 2) may well reflect infants' successful learning during Block 1.

N400 over the right hemisphere during the test phase and the neural activity during the learning phase

As Kaduk (2016) found the N400 effect over right hemisphere electrodes, we specifically analysed how the neural activity during the learning phase can predict the N400 effect prominent over the right-hemisphere electrodes. We constructed a Bayesian LMM of difference in the N400 over the right hemisphere (quantified as an area under the curve on the waveform) between congruent and incongruent conditions on the test phase, as a function of test date, theta-, alpha- and upper-alpha-band power difference between the learning blocks 1 and 2. The hypotheses for this analysis were consistent with the analysis focusing on the N400 over both hemispheres detailed in the previous section.

This model included interaction terms. The model without interaction terms was also constructed to examine the effect given the small sample size of the results. The model fit of the two models were compared using loo function (Vehtari et al., 2020, 2017), which suggested the difference in the model fit between the two models is negligible (elpd difference -0.6, standard error difference 0.7). As the model including interaction terms is more theoretically justified, the model including interaction terms is reported.

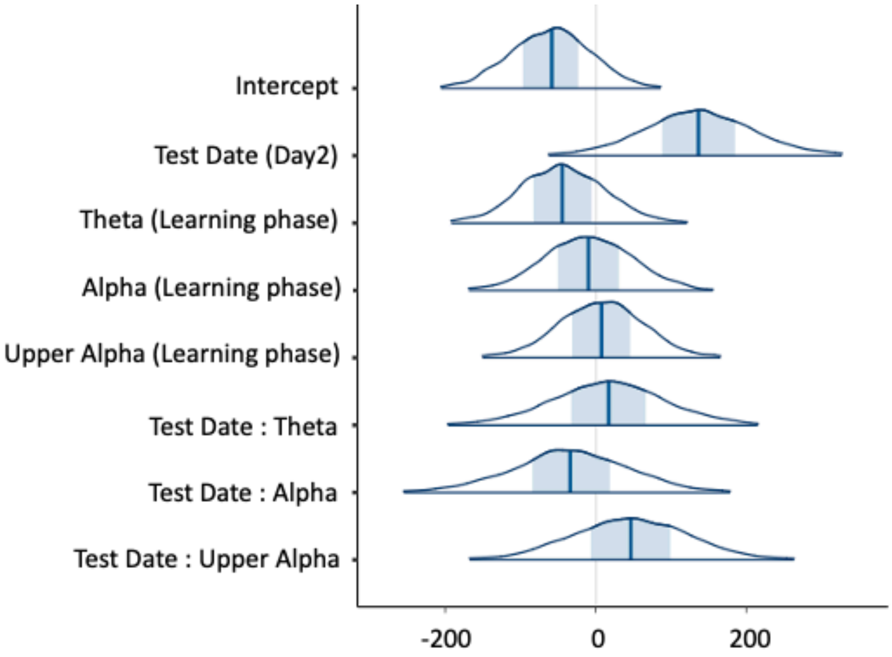
The results of this analysis are summarised in Table 5-15. The posterior distribution of coefficients from this model is shown in Figure 5-30. The visualisation of the main effects and interactions examined are provided in Figures 5-31, 5-32, 5-33, and 5-34. Whilst the results did not strongly support for the interaction between the test dates and oscillatory activity, the main effect of the test date was found to predict the N400 effect over the right hemisphere on the test phase. The difference in the N400 effect quantified as a divergence score (or the area under the curve on the waveform) between the congruent and incongruent conditions on the test phase was larger on the test Day 2 than the test Day 1 (β = 135.79; error = 75.09; 95% credible interval [-14.07; 287.92]). The effect was suggested to be 93% likely to be present (93% credible interval [4.21; 276.52]). Consistent with the previous analysis focusing on the N400 on both hemispheres, this suggests that infants only reliably discriminated the semantic (in)congruency in the presented stimuli on Day 2, not Day 1. This is also consistent with the complementary learning account stressing the important role of offline consolidation in learning (Davis & Gaskell, 2009; McClelland et al., 1995; Norman & O'Reilly, 2003; O'Reilly & Norman, 2002; Weighall et al., 2017).

Table 5-15. Bayesian linear mixed-effects model of the right hemispheric N400 divergence score difference on the test phases as a function of test date, theta-, alpha- and upper-alpha-band power difference between the learning Blocks 1 and 2.

Group-Level	**Effects:**			
Participant	(Number of levels: 14)			
	Estimate	**Est.Error**	**l-95% CI**	**u-95% CI**
sd (Intercept)	57.77	43.33	3.05	159.30

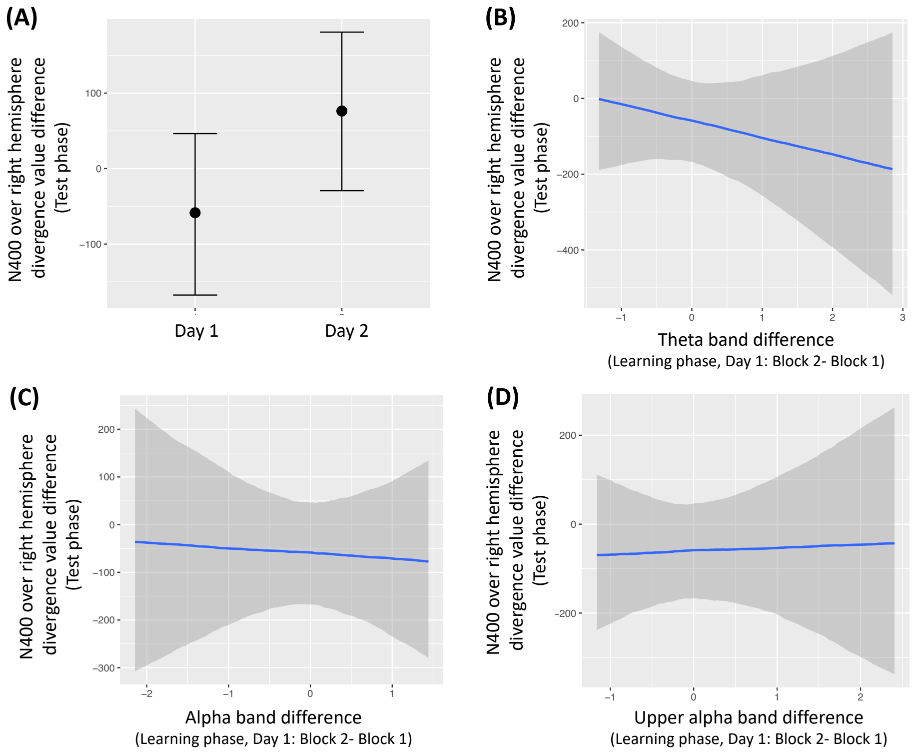
Population-Level Effects:				
	Estimate	**Est.Error**	**l-95% CI**	**u-95% CI**
Intercept	-59.88	54.94	-167.55	46.33
TestDate2	135.79	75.09	-14.07	287.92
ThetaDiff	-43.53	57.67	-155.93	72.42
AlphaDiff	-9.67	60.12	-126.87	107.94
UppAlphaDiff	7.39	58.73	-109.32	123.11
TestDate2 : ThetaDiff	16.23	77.39	-138.29	169.98
TestDate2 : AlphaDiff	-33.96	80.90	-198.87	124.44
TestDate2 : UpperalphaDiff	45.87	79.45	-108.71	200.76

Figure 5-30. Posterior distributions of intercept and coefficients for each predictor.



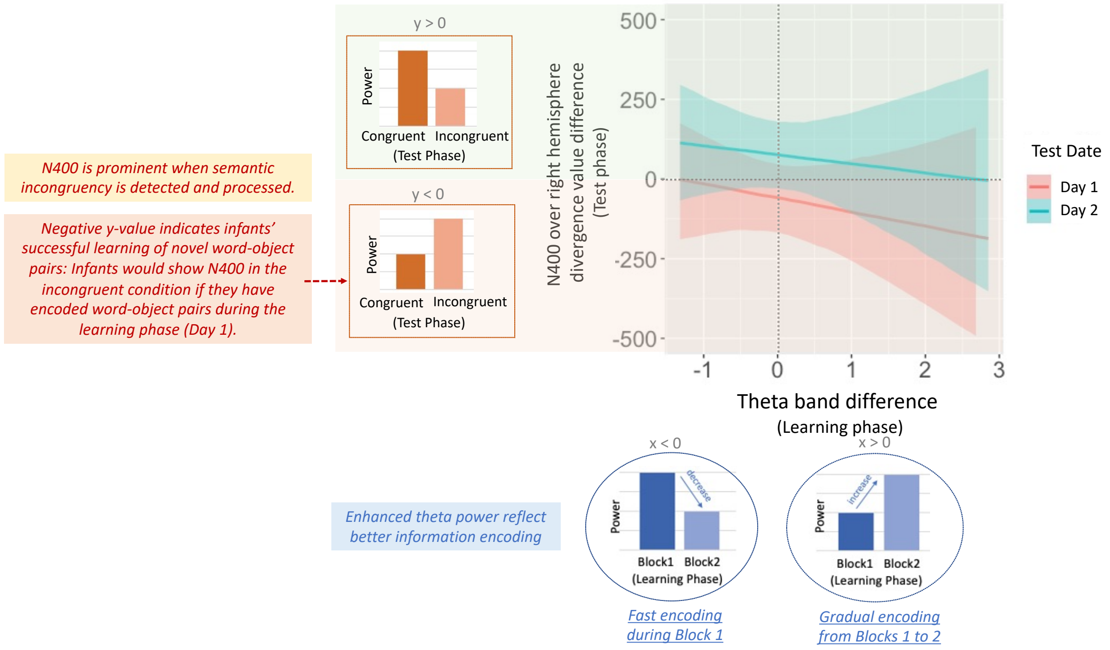
Note. The x-axis shows the estimated coefficients. The y-axis represents the density of the posterior estimates of the intercept and coefficients based on the fitted model. The distributions represent 4000 posterior samples. The central vertical line plotted in the distribution curve shows the estimate mean, and the shaded region represents the 50% probability interval. The distribution tails cover the 99% probability region.

Figure 5-31. Main effect of (A) test date (with or without delay after the learning phase on Day 1), (B) theta-band power theta-band power difference between the learning blocks 1 and 2 (Day 1; M_{Block2}- M_{Block1}), (C) alpha-band power theta-band power difference between the learning blocks 1 and 2 (Day 1; M_{Block2}- M_{Block1}), and (D) upper-alpha-band power theta-band power difference between the learning blocks 1 and 2 (Day 1; M_{Block2} - M_{Block1}) on the N400 over right hemisphere divergence value difference between the congruent and incongruent conditions during the test phase (y-axis: $M_{incongruent}$ - $M_{congruent}$).



Note. The "whisker" in (A) and shaded region in (B), (C), and (D) represent the 95% credible intervals. Each predictor variable (x-axis) is scaled so the mean is 0 and the standard deviation is 1. The outcome variable (y-axis) refers to a difference in N400 over right hemisphere divergence value between the congruent and incongruent conditions of the test phases (Days 1 and 2) (for score calculation, see previous section or Kaduk, 2016). A positive value represents a more prominent N400 during the incongruent condition as compared to the congruent condition. N400 component has been reported upon detection and processing of semantic incongruency (e.g., Friedrich & Friedrici, 2008). Thus, a positive value indicates infants' successful semantic learning. (B) The predictor (x-axis) variable refers to a difference in theta-band power between Blocks 1 and 2 of the learning phase taking place on Day 1. A positive value of predictor (x-axis) means a greater theta-band power during the learning Block 2 than Block 1. A greater theta-band power has been linked to infants' better information encoding (Begus & Bonawitz, 2020), anticipatory attention (Begus et al., 2016; Orekhova et al., 1999), as well as processing of social situations as compared to non-social situations (St. John et al., 2016). A greater theta-band power during Block 1 than 2 (i.e., decrease from Blocks 1 to 2) would indicate infants' fast encoding of the presented information in a social setting, whilst a greater power during Block 2 than 1 (i.e., increase from Blocks 1 to 2) would suggest that infants encoded information better as they heard more labelling by the experimenter. (C) The predictor (x-axis) variable refers to a difference in alpha-band power between Blocks 1 and 2 of the learning phase taking place on Day 1. A positive value represents a greater upper alpha band power during Block 2 than Block 1 (i.e., increase from Blocks 1 to 2). As suppressed alpha-band power has been linked to greater attention allocation (Hoehl et al., 2014; Klimesch, 2012; Michel et al., 2015). Therefore, an increase in alpha-band power from Blocks 1 to 2 indicates infants' attention attenuation during Block 2, suggesting infants' full encoding of the presented information during Block 1. Conversely, a decrease in alpha-band power from Blocks 1 to 2 indicates infants' sustained or increased attention during Block 2 as infants had not yet fully encoded the presented information during Block 1. (D) The predictor (x-axis) variable refers to a difference in upper-alpha-band power between Blocks 1 and 2 of the learning phase taking place on Day 1. A positive value represents a greater upper alpha band power during Block 2 than Block 1 (i.e., an increase from Blocks 1 to 2). As supressed upper-alpha-band activity has been linked to better sustention of semantic knowledge and memory (e.g., Klimsech et al., 1996), we expect that suppressed upper alpha activity should reflect better encoding of semantic information. With the literature proposing the function of neural habituation as a sign of learning (Colombo & Mitchell, 2009; Nelson, 1995; Turk-Browne et al., 2006), an increase in upper-alpha-band power from Blocks 1 to 2 (i.e., habituation from Blocks 1 to 2) may well reflect infants' successful learning during Block 1.

Figure 5-32. Interaction between theta-band power difference between the learning blocks 1 and 2 (x-axis: M_{Block2} - M_{Block1}) and the difference in the N400 divergence scores over the right hemisphere between the congruent and incongruent condition during the test phase (y-axis: $M_{incongruent}$ - $M_{congruent}$).



N400 is prominent when semantic incongruency is detected and processed.

Negative y-value indicates infants' successful learning of novel word-object pairs: Infants would show N400 in the incongruent condition if they have encoded word-object pairs during the learning phase (Day 1).

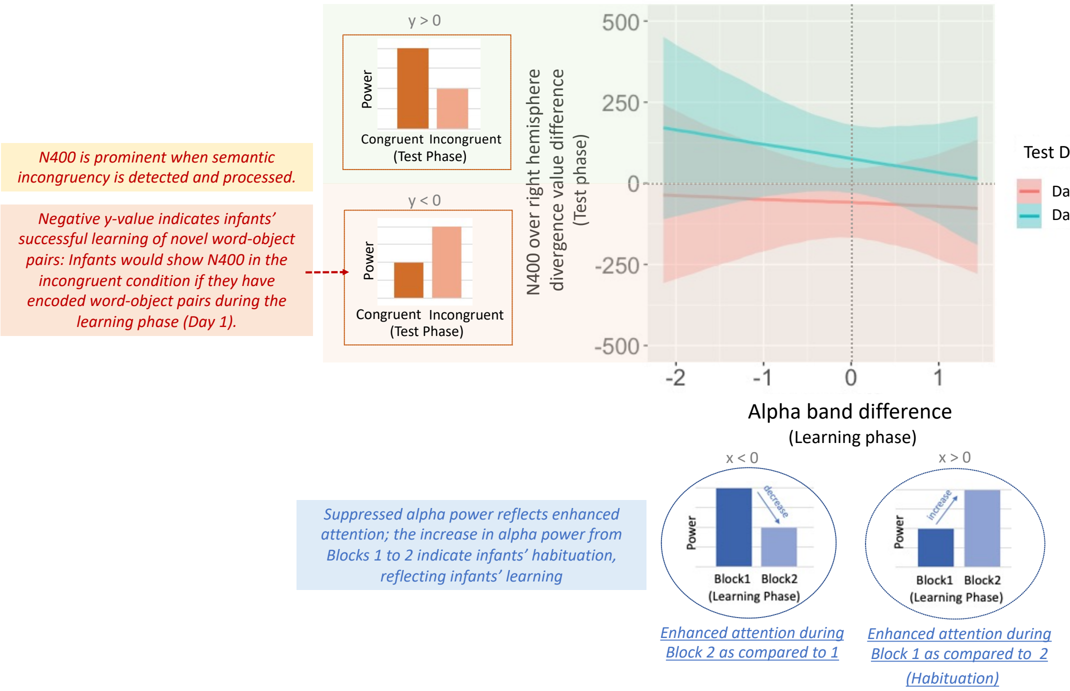
Enhanced theta power reflect better information encoding

Fast encoding during Block 1

Gradual encoding from Blocks 1 to 2

Note. The N400 divergence score reflects the presence or absence of the N400 ERP component over the right hemisphere. The shaded region represents the 95% credible intervals. The predictor variable (x-axis) is scaled so the mean is 0 and the standard deviation is 1. The positive value of predictor (x-axis) means a greater theta-band power during the learning Block 2 than Block 1. A greater theta-band power has been linked to infants' better information encoding (Begus & Bonawitz, 2020), anticipatory attention (Begus et al., 2016; Orekhova et al., 1999), as well as processing of social situations as compared to non-social situations (St. John et al., 2016). A greater theta-band power during Block 1 than 2 (i.e., decrease from Blocks 1 to 2) indicates infants' fast encoding of the presented information in a social setting, whilst a greater power during Block 2 than 1 (i.e., increase from Blocks 1 to 2) suggests that infants encoded information better as they heard more labelling by the experimenter. The outcome variable (y-axis) refers to a difference in N400 over right hemisphere divergence value between the congruent and incongruent conditions of the test phases (Days 1 and 2) (for score calculation, see previous section or Kaduk, 2016). A positive value represents a more prominent N400 during the incongruent condition as compared to the congruent condition. N400 component has been reported upon detection and processing of semantic incongruency (e.g., Friedrich & Friedrici, 2008). Thus, a positive value indicates infants' successful semantic learning.

Figure 5-33. Interaction between alpha-band power difference between the learning blocks 1 and 2 (x-axis: $M_{Block2} - M_{Block1}$*) and the difference in the N400 divergence scores over the right hemisphere between the congruent and incongruent condition during the test phase (y-axis:* $M_{incongruent} - M_{congruent}$*).*



N400 is prominent when semantic incongruency is detected and processed.

Negative y-value indicates infants' successful learning of novel word-object pairs: Infants would show N400 in the incongruent condition if they have encoded word-object pairs during the learning phase (Day 1).

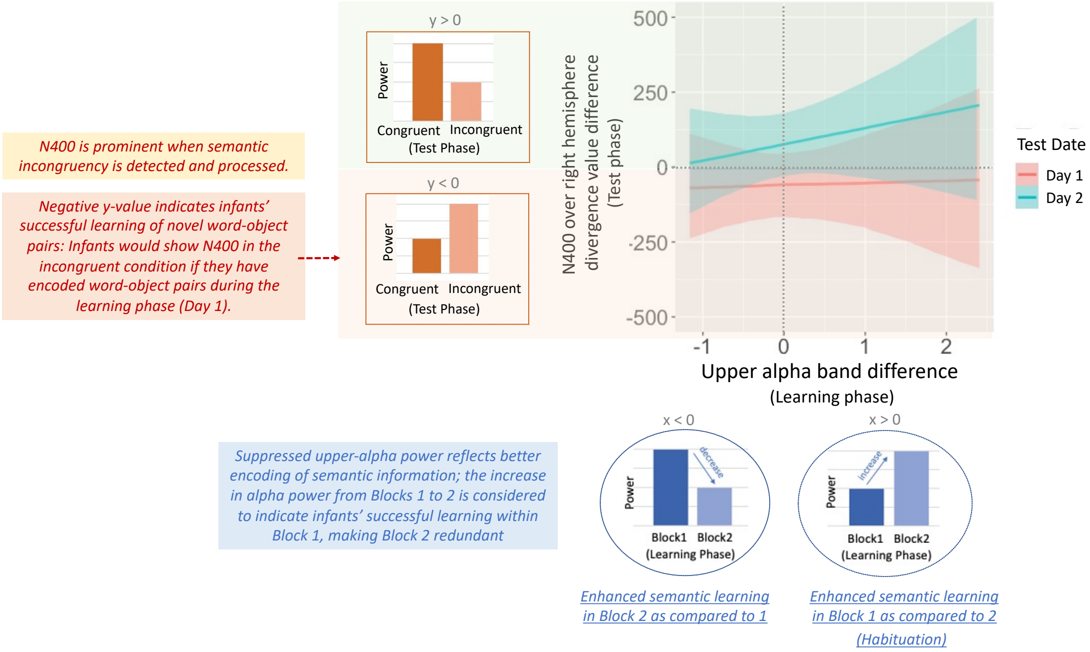
Suppressed alpha power reflects enhanced attention; the increase in alpha power from Blocks 1 to 2 indicate infants' habituation, reflecting infants' learning

Enhanced attention during Block 2 as compared to 1

Enhanced attention during Block 1 as compared to 2 (Habituation)

Note. The N400 divergence score reflects the presence or absence of the N400 ERP component over the right hemisphere. The shaded region represents the 95% credible intervals. The predictor variable (x-axis) is scaled so the mean is 0 and the standard deviation is 1. The positive value of predictor (x-axis) means a greater alpha-band power during the learning Block 2 than Block 1. As suppressed alpha-band power has been linked to greater attention allocation (Hoehl et al., 2014; Klimesch, 2012; Michel et al., 2015). Therefore, an increase in alpha-band power from Blocks 1 to 2 indicates infants' attention attenuation during Block 2, suggesting infants' full encoding of the presented information during Block 1. Conversely, a decrease in alpha-band power from Blocks 1 to 2 indicates infants' sustained or increased attention during Block 2 as infants had not yet fully encoded the presented information during Block 1. The outcome variable (y-axis) refers to a difference in N400 over right hemisphere divergence value between the congruent and incongruent conditions of the test phases (Days 1 and 2) (for score calculation, see previous section or Kaduk, 2016). A positive value represents a more prominent N400 over right hemisphere during the incongruent condition as compared to the congruent condition. N400 component has been reported upon detection and processing of semantic incongruency (e.g., Friedrich & Friedrici, 2008). Thus, a positive value indicates infants' successful semantic learning.

Figure 5-34. Interaction between upper-alpha-band power difference between the learning blocks 1 and 2 (x-axis: M_{Block2} - M_{Block1}) and the difference in the N400 divergence scores over the right hemisphere between the congruent and incongruent condition during the test phase (y-axis: $M_{incongruent}$ - $M_{congruent}$).



N400 is prominent when semantic incongruency is detected and processed.

Negative y-value indicates infants' successful learning of novel word-object pairs: Infants would show N400 in the incongruent condition if they have encoded word-object pairs during the learning phase (Day 1).

Suppressed upper-alpha power reflects better encoding of semantic information; the increase in alpha power from Blocks 1 to 2 is considered to indicate infants' successful learning within Block 1, making Block 2 redundant

Enhanced semantic learning in Block 2 as compared to 1

Enhanced semantic learning in Block 1 as compared to 2 (Habituation)

Note. The N400 divergence score reflects the presence or absence of the N400 ERP component over the right hemisphere. The shaded region represents the 95% credible intervals. The predictor variable (x-axis) is scaled so the mean is 0 and the standard deviation is 1. The positive value of predictor (x-axis) means a greater upper-alpha-band power during the learning Block 2 than Block 1. As supressed upper-alpha-band activity has been linked to better sustention of semantic knowledge and memory (e.g., Klimsech et al., 1996), we expect that suppressed upper alpha activity reflects better encoding of semantic information. With the literature proposing the function of neural habituation as a sign of learning (Colombo & Mitchell, 2009; Nelson, 1995; Turk-Browne et al., 2006), an increase in upper-alpha-band power from Blocks 1 to 2 (i.e., habituation from Blocks 1 to 2) may well reflect infants' successful learning during Block 1. A positive value represents a more prominent N400 over right hemisphere during the incongruent condition as compared to the congruent condition. N400 component has been reported upon detection and processing of semantic incongruency (e.g., Friedrich & Friedrici, 2008). Thus, a positive value indicates infants' successful semantic learning.

Discussion

The current study explored the cognitive processes of 10-month-old infants' word learning that develop during a live social interaction, utilising neural oscillatory measures. We further examined the trajectory of infants' knowledge consolidation that is thought to occur over time (Friedrich et al., 2015; McClelland et al., 1995). We hypothesised that infants' theta oscillatory power would increase and alpha oscillatory power would decrease (i.e., more suppressed) as infants heard more labelling by an experimenter during a live interaction, based on prior studies demonstrating that increased theta-band oscillatory activity can index infants' enhanced information encoding (Begus et al., 2016, 2015; Wass et al., 2018), and supressed alpha-band oscillatory activity reflects attentional allocation as well as the formation of knowledge and memory (Hoehl, Michel, Reid, Parise, & Striano, 2014; Klimesch, 2012; Michel et al., 2015). In this study, we monitored infants' brain activity while they interacted with an experimenter in a naturalistic manner. The experimenter presented them with two novel toys, one at a time, labelling the object with a pseudo-word repeatedly in a naturalistic manner. Each toy was presented three times, so the whole learning phase consisted of six sessions in total. The first, second and third sessions of each object were combined to make three learning blocks. Yet, the examination of video recording data from each learning block suggested that infants were no longer able to maintain their attention to the task after the second block. Therefore, EEG frequency analyses were conducted to examine differences between the first and second blocks of the learning phase. We further assessed the association between the neural activity during the first two blocks of the learning phase and the neural activities that indicated infants' discrimination of correct and incorrect object-label pairs at the subsequent two test phases, with and without a delay. This set of analyses aimed to examine how differences in the neural and cognitive processes during the learning phase might be linked to infants' learning which unfolds over time (i.e., knowledge consolidation).

The neural dynamics during the learning phase

Alpha-band activity as an index of infant learning via neural habituation

Analysing infants' neural activity during the learning phase, we found that alpha-band power (6-9Hz) increased from Blocks 1 to 2 of the learning phase. Supressed alpha-band power has been associated with infants' attention allocation (Hoehl et al., 2014;

Klimesch, 2012; Michel et al., 2015). Hence, this result indicates infants showed enhanced attention during the first block compared to the second block. The attention attenuation between Blocks 1 and 2 can be interpreted as infants' habituation. Infants' habituation to repeated stimuli has been well documented (Colombo & Mitchell, 2009). In infancy research, habituation is typically assessed by looking time, via reduction in looking to a familiar stimulus and subsequent longer looking time towards novel stimuli. Such habituation is typically interpreted as infants' ability to discriminate two stimuli (Malcuit, Pomerleau, & Lamarre, 1988; Sirois & Mareschal, 2002). However, this might also suggest that infants have fully encoded the familiar stimuli. Habituation has been assessed with neural measures in research with adult participants using fMRI. In these studies, neural habituation refers to a reduction in neural activity when stimuli are repeated (Grill-Spector, Henson, & Martin, 2006). Adult neural habituation has been found even when more than one type of stimulus were presented (Turk-Browne et al., 2006), or when test stimuli were presented three days after the familiarisation (Van Turennout, Ellmore, & Martin, 2000). Importantly, neural habituation has been shown to be positively correlated with the degree of effect priming via repetition (Maccotta & Buckner, 2004; Turk-Browne et al., 2006). It is thus suggested that neural habituation reflects the process of encoding long-term perceptual memory (Turk-Browne et al., 2006). Neural habituation in infants has been reported in studies using fNIRS (Benavides-Varela et al., 2011; Bouchon, Nazzi, & Gervain, 2015; Lloyd-Fox et al., 2019; Nakano, Watanabe, Homae, & Taga, 2009), although the cognitive processes underlying such neural habituation in infants is less known compared to that of adults (Lloyd-Fox et al., 2019). Nevertheless, it is possible that infants' neural habituation could also be understood as an index of implicit perceptual memory (Colombo & Mitchell, 2009; Nelson, 1995; Turk-Browne et al., 2006). If this is the case, attention attenuation observed from Blocks 1 to 2 might well be an indication that infants could form word-label associations during Block 1 before going through Block 2. This also means that this temporal dynamics of alpha-band power during a leaning event could be used as a measure of infant novel word learning in a social interaction.

Upper-alpha-band activity as an index of semantic information encoding

Our analysis suggested an increase of upper-alpha-band power (9-10Hz) from Blocks 1 to 2 of the learning phase. Upper-alpha-band suppression (i.e., lower power) in infants has only been examined in few studies (e.g., Kaduk, 2016). Yet, adult studies have indicated that it specifically reflects knowledge consolidation after successful information encoding (Klimesch, Doppelmayr, Pachinger, & Russegger, 1997; Klimesch, Schimke, et al., 1996; Klimesch, 2012). Consistently, infants' upper-alpha-band power has been associated with the N400 ERP component (Kaduk, 2016), which has been used as a reliable marker of infants' semantic knowledge systems in the domains of language (e.g., Friedrich & Friederici, 2006; Kutas & Hillyard, 1980) and other social behaviour (e.g., Kaduk et al, 2016; Reid et al., 2009). Based on these, the upper-alpha-band activity in infants could also be a marker of their semantic understanding. Hence, our results showed that infants' lexical learning observed in the current sample is not merely associative, but reflects the semantic property of word-object associations. Furthermore, the current study extended Friedrich and Friedrici's (2011) findings by demonstrating that infants can encode object-label semantic associations even in a less controlled setting with richer information involving a naturalistic interaction within a very short timeframe, as the more suppressed upper-alpha-band activity was observed during the first block of the learning phase as compared to the later block.

Theta-band activity as an index of infant learning

Despite an increasing number of studies demonstrating infants' theta-band activity as a measure of learning (Begus & Bonawitz, 2020), the current study did not identify any difference in theta-band oscillation across different blocks of the learning phase. It is possible that the degree of infants' information encoding during the live interaction was not robust enough to be detected by theta-band oscillatory activity. The previous study by Kaduk (2016), which analysed infants' object-word recognition following the naturalistic interaction, also failed to find an effect in theta-band activity. Yet, the study reported the effect of learning in terms of the Nc and N400 ERP components as well as upper-alpha-band desynchronisation on the recognition test on the following day. This might also suggest that theta-band activity might not be as sensitive to infants' lexical information encoding in this context as other measures could be.

We believe this has been the first neural evidence demonstrating infants' ability to encode semantic lexical information that occurs in a social interaction within a very short timeframe. Whilst we did not find any reliable effect in theta-band activity indexing infants' novel word learning in a social setting, our findings about the neural activity in alpha- and upper-alpha-band activity jointly suggest that infants' enhanced attention during the initial phase of lexical information exposure in a social setting, as compared to the later phases, might serve as a gateway to successful knowledge consolidation. Importantly, these results indicate that the temporal dynamics of neural activity during a naturalistic social learning event could be used as an index of infant learning.

The neural trajectory from the learning phase to the test phases

To investigate the trajectory of information encoding to knowledge consolidation in the context of novel word learning, we examined how the theta-band, alpha-band and upper-alpha-band oscillatory activity during the learning phase might predict infants' recognition of word-label associations assessed after the learning phase with and without delay. The question was addressed using Bayesian linear mixed-effects models of the neural activity of interest during the test phase as a function of test date as well as the neural activity during the learning phase. As prior studies suggested that infants' knowledge consolidation occurs with delay, during which offline consolidation occurs (Friedrich & Friederici, 2011; Kaduk, 2016), we expected that the degree to which the neural activity during the learning phase (taking place on Day 1) predict the neural activity during the test phase would differ between the Day 1 test phase (which took place immediately after the learning phase) and Day 2 test phase (which took place one day after the learning phase). This examination was expected to enable us to investigate whether the dynamics of neural activity in upper-alpha and alpha frequency bands is associated with infants' successful learning immediately after and a day after the learning event. Furthermore, it was hoped to also enable us to explore potential individual differences in infants' cognitive processes during the learning event which might modulate the subsequent offline consolidation, as there has been mixed evidence with regards to whether infants younger than 14 months of age could retain fast-mapped information for longer than a day (Friedrich & Friederici, 2005b, 2008; Friedrich et al., 2015).

Examining the relationship between the neural activity during the learning phase and the neural correlates of infants' semantic knowledge, the current study found that infants' ability to encode semantic information as well as attentional skills informed how novel words are encoded and stored as knowledge from a learning event involving naturalistic social interactions.

Individual differences in semantic fast-encoding ability in a social setting

Our previous analysis of the neural activity during the learning phase indicated that an increase in upper-alpha-band power from Blocks 1 to 2 of the learning phase suggest infants' successful semantic word learning that occurs within the first few minutes of information presentation in a social setting (Klimesch et al., 1997). Our Bayesian LMMs analysis further found that this increase has been linked to a neural index of successful discrimination of unexpected events as opposed to expected events, reflected in a theta-band power burst that occurs when detecting and processing unexpected events (Köster et al., 2019, 2021). That is, on the Day 1 test phase, infants who showed an increase in upper-alpha-band activity from Blocks 1 to 2 on the learning phase exhibited a greater theta power for the incongruent condition than the congruent condition. Conversely, infants who showed a decrease in upper-alpha-band activity from Blocks 1 to 2 on the learning phase showed a greater theta power for the congruent condition than the incongruent condition. Taken together, these results support that the neural dynamics of upper-alpha-band activity during a social learning event can successfully be used as an index of infants' learning, in such a way that an increase in upper-alpha-band power may well reflect infants' successful learning of semantic information.

Importantly, this also suggests that there are individual differences in the ability to encode and consolidate semantic knowledge in a social interaction, in such a way that some infants may be faster than others in encoding object-label semantic associations in a social setting involving rich and complex information. Such individual differences could explain the null result reported by Kaduk (2016) that they did not find any effect in theta oscillatory power on either test date. Perhaps, a 6-minute social interaction phase was enough for some infants but not for others to encode word-object associations, and those who did not fully encode the information during the learning event continued to encode the information during the test phase that occurred immediately after the interaction. It is important to note that the direction of the effect was reversed on the Day 2 test day where the memory retention of 24 hours was assessed. This suggests that infants who take longer to encode semantic information on the first day do benefit from offline consolidation, as they can show the signs of semantic knowledge on the following day of the learning phase.

Our analysis further indicated that the dynamics of upper-alpha-band activity during the learng event was associated with the difference in the amplitude of Nc ERP component between congruent and incongruent word-object pairs presented at the test phases after the learning event with and withour a delay. More specifically, infants who showed an increase in upper-alpha-band power from Blocks 1 to 2 during the learning phase did not show any difference in the Nc mean amplitude when they were shown the congruent and incongruent word-object pairs immediately after the learning event, whist infants who showed a decrease in upper-alpha-band power from Blocks 1 to 2 did show a difference in the Nc mean amplitude immediately after the learning event. The Nc ERP component has been linked to infants' attention (Richards, 2003). Given that an increase in upper-alpha-band power from Blocks 1 to 2 of the learning phase can index infants' semantic encoding during the first block (see previous section), no significant difference in the Nc amplitude in infants who shoed such an alpha-band power increase from Blocks 1 to 2 may well be habituation, or attentional disengagement (i.e., randomly allocated attention) which occurs after they have successfully encoded the target information (Colombo & Mitchell, 2009; Nelson, 1995; Turk-Browne et al., 2006). On the other hand, for infants who were still learning the semantic association of word-object pairs during the Block 2 continuing from the Block 1 of the learning phase (reflected in a decreased upper-alpha-band power during Block 2 compared to Block 1), the attentional engagement during the immediate test phase should still be high (i.e., not habituated yet) (Colombo & Mitchell, 2009), leading to the differential amplitude of the Nc between congruent and incongruent object-label pairs presented at the recognition test. Taken together, these pieces of evidence regarding the link between upper alpha band activity during the learning event and the Nc amplitude at the test phases suggests the potential use of upper alpha band activity as an index of infants' semantic learning.

It was also found that individual differences in establishing the semantic knowledge system in a short time frame in a social setting, as reflected in an increase in upper-alpha-band power from Blocks 1 to 2 of the learning event, modulated the degree to which infants recalled the encoded information on a delayed test phase. On the delayed test phase, infants who encoded and consolidated semantic information faster (i.e., an increase in upper-alpha-band power from Blocks 1 to 2 of the learning phase) showed a larger difference in the Nc amplitude, compared to those who were still learning the information beyond Block 1 of the learning event (i.e., a decrease in upper-alpha-band power from Blocks 1 to 2 of the learning phase). This indicates that individual differences in infants' ability to learn words in a social situation influenced the degree of knowledge consolidation, in such a way that infants who can extract and encode semantic lexical information in a complex social setting within a shorter timeframe than others can build a more robust memory structure after the offline consolidation (Csibra & Gergely, 2009; Kuhl, 2007).

Individual differences in attentional engagement as a scaffold to consolidate semantic knowledge

Our analysis also found that the neural activity in the alpha band during the learning phase was reliably associated with the neural activity in upper-alpha band during the test phases. That is, infants who showed an increase in alpha-band power from Blocks 1 to 2 of the learning phase exhibited a more suppressed power for the congruent condition than the incongruent condition of the test phase which took place immediately after the learning event. This pattern of upper-alpha-band activity during the test phase has been found to index infants' successful word learning (Kaduk, 2016). Therefore, this demonstrates that an increase in alpha-band power during a learning event does reflect infants' successful learning of lexical semantic information (i.e., words), supporting of our interpretation of our results regarding the alpha-band neural activity during the learning event.

Importantly, this result also indicates that individual differences in attentional engagement in the learning phase involving a social interaction modulated the trajectory of semantic knowledge consolidation, as the direction of this association differed between the test phases on the Day 1 and Day 2. As mentioned, infants who showed an increase in alpha-band power from Blocks 1 to 2 of the learning phase (i.e., suppression during the Block 1) exhibited a supressed upper-alpha band power in the congruent condition as compared to the incongruent condition immediately after the learning event as well as after a delay of 24h (i.e., both Day 1 and Day 2 test), reflecting successful semantic information encoding and consolidation (Kaduk, 2016). Yet, infants who showed a decrease in alpha-band power from Blocks 1 to 2 of the learning phase only showed suppressed upper-alpha-band activity for the congruent condition, contrasted with the incongruent condition, on the delayed test phase (i.e., Day 2 test), but not immediately after the learning phase (i.e., Day 1 test). These indicate the critical role that infants' attentional skills play in the context of infants' word learning previous literature (Kuhl, 2007; Posner, 2004; Smith & Yu, 2012; Yoshida, Pons, Maye, & Werker, 2010; Yu & Smith, 2010, 2012). This also aligns well with a proposal put forward by Kuhl and colleagues (Kuhl, 2007a; Kuhl et al., 2008), which posited that social interaction is "essential for natural speech leaning" (p. 110, Kuhl, 2007). According to their view, attention attracted to social agents as well as arousal induced by social interaction contribute to an overall increase in the quantity and quality of language information that infants code and remember (Kuhl, 2007a; Posner, 2004). In line with this account, attention has been shown to play an important role in learning language, as 10-month-old ‘high-attender’ infants were able to learn phonetics of their non-native language from phonetically bimodal stimulus, whilst ‘low-attender’ infants did not (Yoshida et al., 2010). Other studies have also indicated the effect of individual differences in attention on infant learning (Smith & Yu, 2012; Yu & Smith, 2010). Furthermore, the association between social attention and word learning has been reported in a behavioural study with older children aged 20 months (Shneidman, Buresh, Shimpi, Knight-Schwarz, & Woodward, 2009). Adding to the existing evidence, the current study has provided the first neural evidence that infants’ social attentional skills and language development are intertwined.

Upper-alpha oscillatory activity as a measure of infants’ semantic knowledge systems

Given the mixed evidence from utilising the N400 ERP component as a measure of infant semantic knowledge consolidation, it is important to discuss the use of this component as well as to explore an alternative measure. Kaduk (2016) reported the N400 effect indexing infants’ semantic knowledge consolidation on the delayed test phase but not on the immediate test phase. Supporting this, the results of the current study indicated that the N400 effect was stronger on Day 2 than Day 1. The current study further reported that the N400 effect on the test phases was not significantly modulated by the oscillatory activity in theta-, alpha- and upper-alpha frequency bands during the learning phase. Friedrich and Friedrici (2011, 2017) have argued that the N400 component only represents a strong form of semantic memory structure, as it can be missing when other ERP components are present that suggest the formation of word-object associations, such as N200-500 (Friedrich & Friederici, 2011, 2017). Given that upper-alpha band has been linked to the N400 (Kaduk, 2016), upper-alpha-band oscillatory activity might be a more sensitive measure of infants’ semantic information encoding and consolidation, as it can index individual differences as shown in the current study. Future studies could utilise both measures so it can assess a robust memory structure via N400, as well as a less robust form of semantic knowledge systems via upper-alpha-band oscillatory activity. Adding this frequency measure, it would be possible to investigate further about the mixed evidence in the current literature, with regards to whether infants younger than 14 months of age could retain fast-mapped information overnight, making use of offline consolidation (Friedrich et al., 2015; James et al., 2017).

Conclusion and future directions

To sum, these findings jointly demonstrate that infants are able to encode semantic word-object associations within a few minutes of information exposure even in a less controlled naturalistic interactional setting. Critically, our findings suggested that individual differences in such a fast-encoding ability in a social situation influence the trajectory of knowledge consolidation. Attentional engagement during the learning phase was also found to be a critical factor for infants' word learning and the subsequent successful knowledge consolidation.

Evidence from the current study has taken our understanding further with regards to how infants learn novel words from a continuous naturalistic speech. New questions arising from this study include which factors of a social interaction (e.g., direct eye contact, joint attention, infant-directed speech) facilitate infants' attentional engagement and semantic lexical learning. Such a question could not be addressed by the design of the study presented in this work, where an experimenter demonstrated the task in as naturalistic a way as possible, using various ostensive signals. Yet, we believe that our study provided a reliable paradigm and empirical data to drive hypotheses for further investigation to examine such a question. Moreover, the methods used in the current study of exploring the link between social attention and learning outcomes could be applied for research investigating the neural mechanisms by which certain atypically developing population, such as children with autism, could develop both difficulties in social and language domains (Chita-Tegmark, 2016; Geraldine Dawson, Bernier, & Ring, 2012; Eigsti, De Marchena, Schuh, & Kelley, 2011). We believe this study has laid a foundation for future investigation to connect how the functions of the 'social brain' at an earlier age influence the later development in various domains including, but not limited to, cognition, social cognition and language.

Acknowledgement

The study reported here is a secondary data analysis, which I conducted during the COVID-19 pandemic in 2020 and 2021 when in-person data collection was not possible. I thank Dr Katharina Kaduk for giving me the permission to analyse the data she previously collected with her colleagues. The original study was designed and conducted by Dr Katharina Kaduk, Dr Kirsty Dunn, Dr Eugenio Parise, Professor Stefanie Hoehl and Professor Vincent Reid. My analysis focused on the portion of their data which had not been analysed for the original study, which was included in a doctoral thesis of Dr Kaduk. I extend my thanks to Barrie Usherwood for his support with MATLAB coding, Chiara Capparini and Maddy Anderson for the validation of behavioural coding, and Michaela DeBolt and Dr Samuel David Jones for their support with the R scripts for linear mixed-effects models.

Author Contribution

Sayaka Kidby: analysis design, analysis, writing, review. Katharina Kaduk: data collection. Kirsty Dunn: data collection, analysis design, review. Vincent Reid: analysis design, review.

Chapter 6

Towards the understanding of the neural mechanisms of Piagetian A-not-B search error

Chapter Introduction

In this chapter, we provide another example of infant cognition research taking the second-person cognitive neuroscience approach, in which infant brain activity was monitored while the experimenter and the child engaged in a naturalistic and reciprocal social interaction. In the existing studies, even when there is an interactional element in experimental paradigms (e.g., Hoehl, Michel, et al., 2014; Parise et al., 2008; Striano et al., 2006), the behaviour of an adult who interacts with an infant is often relatively restricted, limiting the level of reciprocity and contingency of the interaction that emerges. Here, we aimed to evaluate the feasibility of research where the constraints on the reciprocity and contingency of interaction are loosened.

Here, we implemented electrophysiological measures in a traditional Piagetian A-not-B search task. Typically, this task involves an experimenter and an infant playing a 'hide-and find' game. It is a common practice that an experimenter engages with an infant in a naturalistic manner so they are encouraged to find a hidden toy. These social exchanges between an experimenter and an infant are a critical feature of the task (Dunn & Bremner, 2019; Topál et al., 2008). The presented study here has shown that it is feasible to utilise electrophysiological measures without sacrificing the social nature of the task.

Abstract

Despite the repeated replication of the infant perseverative error ("A-not-B error") originally reported by Piaget (1954), no existing accounts fully explain all the reports thus far regarding the settings in which this error occurs, leaving the question open as to why the error occurs and what aspect of infants' cognition that the A-not-B error reflects. As different measures have often yielded mixed evidence, the use of multiple measures is critical to investigate the cognitive processes underlying this search error (Dunn & Bremner, 2019). The use of neural measures indexing online cognitive processes is promising to shed light on the matter. The current study monitored infants' neural activity using EEG whilst they engaged in the traditional version of the A-not-B search task, expecting to find systematic neural differences between accurate and inaccurate searches of 9-month-old infants (*N*=35). Theta-band oscillatory activity was expected to index different search performances, as it has been linked to information processing in social (St. John et al., 2016) and non-social contexts (Begus et al., 2015), as well as attention based on anticipation (Orekhova, Stroganova, & Posikera, 1999; Orekhova, Stroganova, Posikera, & Elam, 2006). The behavioural analysis indicated that the sample of this study did not replicate the typical A-not-B error, likely due to the setup. This limited the inferences that could be made from the analysis of neural measures in relation to A-not-B errors. Yet, a methodological procedure and analysis pipeline for future EEG research using a live, traditional Piagetian A-not-B search task has been established. The current study showed the feasibility of exploring the cognitive processes that underlie behavioural tasks using electrophysiological measures. Importantly, the implementation of neural measures can be done in such a way where it does not interfere with conventional behavioural task procedures such as this traditional Piagetian A-not-B search task.

Introduction

The A-not-B search error, or perseverative error, was first reported by Piaget (1954). The error occurs in a hide-and-find task, which typically involves two hiding locations. Infants of 9 months of age observe an adult hide an object in one location (A). Following a delay, they are encouraged to search for and retrieve the object whereby they are typically successful. After a number of successful 'A trials', the adult hides the object in the second location (B). Infants can see the location swich, but they often fail to retrieve the object after the switch ('B trials'), as they persist in searching at the previous A location. There have been mixed reports regarding the developmental trajectory of this error across the first postnatal year of life. Some have documented the improvement in search performance as children age (Diamond, 1985; Gratch & Landers, 1971), whereas others have provided empirical and anecdotal reports of a U-shaped trajectory that the error only occurs from 7 to 12 months of age, but not earlier (Bremner, 1998; Clearfield, Diedrich, Smith, & Thelen, 2006).

Whilst this perseverative error has been amongst the most repeatedly studied and replicated findings in the developmental psychology, there has been little consensus with regards to why the error occurs. There has been a continuing controversy regarding whether search performance reflects the development of object permanence (Piaget, 1954). Object permanence refers to an understanding that an object continues to exist at the same location even when it is out of sight (Piaget, 1954, 1966). According to Piaget's view, infants' understanding of object permanence is still immature and hence they would believe that an object would exist independently of their own reaches, because they egocentrically associate the appearance of an object with their searching at a particular location. Thus, it was argued that the error should be a reflection of infants' limited and fragile concept of object permanence at around the age of 9 months (Piaget, 1954).

Conversely, studies using a violation of expectation paradigm with preferential looking measures suggest that a concept of object permanence should already be robust at 9 months of age (e.g., Baillargeon, 1987; Baillargeon, Devos, & Graber, 1989; Baillargeon & Graber, 1988). A violation of expectation paradigm typically presents infants with unexpected events as compared to expected events regarding physical world principals, and longer looking time towards unexpected events is typically interpreted as evidence for infants' knowledge of the underlying principal (e.g. Baillargeon, 1987). Studies utilising this paradigm have demonstrated that infants as young as 3.5 months of age understand object permanence (Baillargeon, 1987). For instance, Baillargeon (1987) presented 5-month-old infants with two sets of “drawbridge” events, both of which included a screen and a box placed behind the screen. In one scenario (‘expected’), infants saw a screen rotating towards the box, occluding the box, and stopping when it hit the box. In the other scenario (‘unexpected’), a screen rotate towards the box, occluding the box, but continuing to move through 180 degrees, appearing to pass through the space where the box was (Baillargeon, 1987). In this study, infants reliably looked longer at the ‘unexpected’ event, indicating that they expected the box to exist even when it is not visible to infants occluded by the screen, rather than disappear and let the screen pass through. This suggests that infants aged 3.5 months do have object permanence and understand that the presence of an object does not rely on their direct perceptual input. In the context of the A-not-B error, infants’ object permanence has been investigated using a preferential looking measure regarding the situation with two hiding locations, resembling the original A-not-B search task reported by Piaget (1954) but without actual reaching (Baillargeon, Devos, & Graber, 1989; Baillargeon & Graber, 1988). These studies have consistently reported evidence that infants aged 9 months and younger have an understanding of object permanence, as infants show longer looking when an object is retrieved from a location where it should not be (Baillargeon, Devos, & Graber, 1989; Baillargeon & Graber, 1988).

As the perseverative error occurs in 9-month-old infants who have been shown to have a stable concept of object permanence, it has even been suggested that this A-not-B error might not reflect infants’ understanding of object permanence at all (Smith, Thelen, Titzer, & McLin, 1999). Smith and colleagues (1999) investigated whether infants err even when no object is hidden underneath the cover. In their study, two lids were placed over two empty wells (i.e., hiding locations). For A trials, an experimenter waved a lid (i.e., cover) placed over location A (‘A lid’) to attract the infant’s attention, whilst the B lid remained over the location B. For B trials, the experimenter waved the B lid, whilst the A lid remained over the location A. Even in such a “lids-only” task, infants still displayed the perseverative reach towards the B location, making the typical A-not-B error. It was argued that if the error is linked to object permanence (or the presence of an object), the A-not-B search task with no hidden object should not lead to the A-not-B error. Yet, an alternative explanation is possible. Perhaps, infants see a lid as an object to be retrieved, rather than an object to hide something else. If so, a lid acts as an object for infants to search for in this ‘lids-only’ version of the search task. Hence, this evidence does not fully contradict with the idea that the perseverative error is linked to object permanence. Other studies have also provided evidence to suggest that the presence of an object is key to this A-not-B search error. For instance, Munakata (1997) examined whether 10-month-old infants can differentiate covers with and without an object underneath. On the A trials, only a cover was used without a toy, and infants were encouraged to reach towards the A location. On the B trials, when only a cover was used, infants typically made the A-not-B error. On the contrary, when the toy was introduced on the B trials and hidden beneath the cover, infants did not make the A-not-B error. Furthermore, when the toy was used during the A trials, regardless of the conditions for the B trials (i.e., cover-only or hidden-toy), infants made the A-not-B error on the B trials. The difference between these cover-only and hidden-toy conditions were also examined by Bremner and Bryant (2001), who offered supporting evidence that infants respond differently to the search task depending on whether there is an object hidden beneath the cover or not. These suggest that the presence of an object does play a role in this perseverative error.

As the evidence suggests that infants’ understanding of object permanence underlie the search errors, we must consider what might contribute to the currently mixed evidence regarding the age at which evidence for object permanence becomes reliable. It has been suggested that the choice of research measures may explain the disparity (Dunn & Bremner, 2019). Diamond (1988) suggested interpretations of looking and reaching do not always correspond to each other, as infants often look to the accurate location (location B) on the B trial, although they subsequently reach towards the wrong location (location A). It was argued that infants do understand that the object is hidden in the location B but do not have the capacity to control their reaching behaviour based on the knowledge (Diamond, 1988). Supporting Diamond’s idea, Hofstadter and Reznick (1996) also demonstrated that when the directions of infants’ gaze and reach were different, the direction of gaze is more likely to be correct (Hofstadter & Reznick, 1996). This suggests that the A-not-B search error might involve other cognitive skills and processes beyond object permanence, such as motor inhibitory control, and looking and reaching measures might be highlighting different processes underlying infant behaviour at the A-not-B search task.

Indeed, a number of other explanations have been put forward for this search error (Dunn & Bremner, 2019). For example, evidence suggests that a delay between when the object is hidden and when the infant is allowed to search plays a critical role in infants' perseverative error (Diamond, 1985; Gratch, Appel, Evans, LeCompte, & Wright, 1974; Harris, 1973; Wellman, Cross, Bartsch, & Harris, 1986). Gratch and colleagues (1974) examined how the duration of this delay affect infants' search performance. They found that infants do not make search errors on B trials if there is no delay. If there is a 1-second delay, however, they do make an error (Gratch et al., 1974). Other studies also reported that infants who do not experience a delay perform significantly better than those who experience a 5-second delay (Diamond, 1985; Harris, 1973). At the same time, infants are typically able to retrieve the object on the A trials even if there is a delay period of as long as 7 seconds before they are allowed to reach (Gratch et al., 1974), suggesting that the delay only affects infant behaviour after the hiding location switch. Given these findings, it has been proposed that the perseverative error should result from a combination of infants' limited memory and inhibitory skills (Diamond, 1985). According to this view, infants need to be able to hold a memory for the location B throughout the delay period, and then use this information to hinder a stronger motor memory trace established through the repeated reaching to the location A on multiple A trials (Diamond, 1985). When there is a longer delay period, it contributes to a decay in memory about the location B. This makes it harder for this location memory to override the motor memory trace of reaching towards the location A. Therefore, when infants do not have to hold the memory during a delay period, it prevents the memory from decaying and infants should not err (Diamond, 1985). This account can explain the mixed evidence between studies using looking and reaching measures, as looking measures do not require infants' inhibitory skills against the memory trace of reaching movement towards the location A. However, the procedure used in a study which reported infants' improved search performance at a no-delay procedure involved multiple reversals of A and B locations, whereby the hiding location was switched back to the A location after the B trials (Diamond, 1985). Infants' performance in such a 'reversed' procedure has been shown to be poorer compared to a standard procedure only involving one location switch, likely due to the increased task demand (Dunn & Bremner, 2019). Thus, the improvement in infant search performance reported in such a multiple reversal task by removing a delay period might be rather related to the reduced task demand in general, not just the memory demand.

A significant challenge to the memory account is found in evidence showing that infants still make the A-not-B error when the object is not hidden. An example of this non-hidden-object procedure is seen in the aforementioned lids-only study by Smith and colleagues (1999). If infants see the lid as an object to be searched for and retrieved, the lid could be interpreted as a non-hidden object. Indeed, infants do make the A-not-B error in this version of the task (Smith et al., 1999). Other researchers have used a transparent cover to examine the memory account (Bremner & Knowles, 1984; Butterworth, 1977). With a transparent cover, the object is still visible even when the cover is placed. Butterworth (1977) introduced a transparent cover into the task, and analysed the consistency of the 'error run'. The error run refers to the number of consecutive errors made before an accurate search. The analysis of this index found that infants consistently reached for the incorrect location during the B trials even when the transparent cover was used and the object was visible. This indicates that infants' memory decay cannot fully explain the error. If the object is in view under a transparent cover, infants do not need to remember where the object is hidden. Nonetheless, they do make the perseverative errors.

Along with the memory decay, the influence of motor bias on infant behaviour was proposed as an important element of infants' poor search performance, which is thought to develop over the course of the A trials (Diamond, 1985; Smith et al., 1999; Thelen, Schöner, Scheier, & Smith, 2001). In line with this, evidence has shown that previous experiences of reaching to the B location once or twice improve infants search performance on the B trials (Smith et al., 1999). This suggests that motor bias from previous reaching movement affects the likelihood of perseverative error. Looking to the developmental trajectory of reaching, Clearfield and colleagues (2006) reported that infants at 7 months of age and older made the A-not-B error, whereas 5-month-old infants did not. To explain this developmental change, the authors examined the kinematics of infants' reaching movement, and demonstrated that infants' smooth and straight movements towards the B location only emerge at around 7- to 8-months of age, while younger infants' reach on B trials is unsteadily controlled. These suggest that young infants' reaches are too variable and poorly controlled to build a strong motor memory, leading to the 'improved' search in young infants as compared to older infants. This explains the U-shaped developmental trajectory of this perseverative error, in that the error occurs from around 9 months of age until 12 months of age but not earlier (Bremner, 1998; Clearfield et al., 2006). To further support the motor bias account, evidence suggests that infants are more likely to succeed when they can simply repeat the same reach to find the object (Bremner & Bryant, 1977; Butterworth, 1975). For instance, Butterworth (1975) used a movable platform which held two containers to hide an object. There were two groups of infants. For both groups, an object was hidden in the central location, just in front of the infant. When these A trials finished, the location platform was moved so that the central location was then to the left of the infant. On the B trials, the first group of infants saw an object being hidden in the central location, which was previously to the right of the infant during the A trials (i.e., the container in which the object was hidden changed, whereas the relative position of the hiding location to the infant remained the same). The other group of infants saw an object being hidden in the left-hand side position, which used to be the A location (i.e., the container in which the object was hidden remained the same, whereas the relative position of the container to the infant changed). It was reported that when infants can simply repeat the previous reach from the A location (i.e., the hiding location stays the same relative to the infant), they were more likely to succeed in retrieving the object.

Even without reaching, a “motor bias” has been shown to be formed through observations, suggesting infants’ processing of the experimenter’s behaviour of demonstrating the hide-and-find task also plays a role in the occurrence of the perseverative error (Landers, 1971). Landers (1971) compared infants’ search performance at the B trials when, during the A trials, they were encouraged to reach to a hiding location to retrieve a hidden object, and when they only observed the experimenter searching and retrieving the object. No difference was found in terms of the search performance, suggesting that observations are sufficient to create the motor bias for infants to make the A-not-B error, whilst infants’ reaching towards the A location led to longer error runs (i.e., longer consecutive errors) during the B trials (Landers, 1971). The effect of observation on infants’ perseverative error is likely to be due to the mirror neuron networks that are known to be activated by action observation and execution (Paulus, Hunnius, Vissers, & Bekkering, 2011; Saby, Meltzoff, & Marshall, 2013; Southgate, Johnson, Karoui, & Csibra, 2010; van Elk, van Schie, Hunnius, Vesper, & Bekkering, 2008). The evidence provided by Longo and Bertenthal (2006) further supports the role of the mirror neuron systems at play in the observation version of the A-not-B search task. Motor representation is closely linked to their own motor experiences (Cannon, Woodward, Gredebäck, von Hofsten, & Turek, 2012; Daum, Prinz, & Aschersleben, 2011; Loucks & Sommerville, 2012; Sommerville, Woodward, & Needham, 2005). Therefore, observing the action that infants cannot produce by themselves should not activate the mirror neuron systems as strongly as the observation of the action in their own repertoire would. Indeed, Longo and Bertenthal (2006) showed that infants who observed an experimenter making an ipsilateral reach (i.e., reaching to the same side) during the A trials were more likely to make the A-not-B error than those who observed an experimenter making a contralateral reach (i.e., reaching across body). Taken together, infants' interactions with the experimenter who demonstrates the hide-and-find task are likely to play a role in the formation and strengthening of the motor trace, which contributes to infants' perseverative error.

Perceptual aspects of the setup other than the examiner's behaviour have also been reported as influential factors of infant search performance. A meta-analysis reviewing 30 studies investigating the A-not-B error reported that the number of hiding locations impact infants' search behaviour, as well as infant age and the duration of the delay period (Wellman et al., 1986). Interestingly though, the direction of the effect of the number of hiding locations was found to be inconsistent (Wellman et al., 1986). The authors of this meta-analysis suggested that it might be a combination of delay and hiding locations that contributes to the search error, whilst a specific combination that leads to a successful or false search was not identified (Wellman et al., 1986). Yet, a close examination of studies using multiple hiding locations suggests that infants do not simply repeat the same reach when the hiding location changes. For instance, Cummings and Bjork (1981) examined infant search behaviour using five hiding locations. The A (first hiding location) and B locations (the second hiding location) were positioned on the either end of the apparatus, so the distance between the A and B locations was maximal. It was reported that infants rarely searched at the A location during the B trials. As they repeated the B trials, their search responses became progressively closer to the correct hiding location. Moreover, the errors made on the B trials were clustered around the correct B location, rather the A location. In a follow-up study by the same authors, infants were presented with five locations, and were introduced to a new hiding location (the location C for C trials) (Cummings & Bjork, 1983). It was again demonstrated that infants' incorrect search on the B and C trials tended to cluster around the correct object location (Cummings & Bjork, 1983). These studies suggest that infants are able to encode, store and retrieve information regarding the object location, but such information is not fully used to guide their next action. Therefore, the perseverative error must result from an interaction among infants' information processing, memory as well as motor planning, control, and execution.

Given that the existing evidence suggests a multitude of factors which influence infants' search behaviour at the A-not-B search task (Marcovitch & Zelazo, 1999; Wellman et al., 1986), the dynamic systems model provides a more comprehensive account of infants' A-not-B error instead of searching for a single cause (Dineva & Schöner, 2018; Smith & Thelen, 1994; Smith et al., 1999; Thelen et al., 2001). This model takes the embodied cognition approach, and explains infant behaviour as an outcome of bodily interactions with the surrounding environment. This theory considers multiple inputs including 'task input' which is consistently present throughout the task (e.g., object, cover, hiding locations) and temporary 'specific input' (e.g., experimenter's behaviour). These inputs interact with infants' memory and abilities to plan, control, and execute reaching behaviour. The theory explains the occurrence of the A-not-B error in the neural activation field, where neurons 'vote' for either direction to reach, based on the input. These neural activation patterns create a 'peak', and when the reaching decision is 'read out' at a certain time point (e.g., at the end of the delay), a direction containing a peak larger than certain threshold will be the location where infants reach. There is also 'memory input', which remembers previous decisions to reach towards the location A or B. This input accounts for a motor trace that has been repeatedly shown to be influential to infants' behaviour at the search task (Diamond, 1985; Smith et al., 1999; Thelen, Schöner, Scheier, & Smith, 2001), whilst also taking into account a spontaneous reach that infants could make (Dineva & Schöner, 2018). This model can provide an explanation for the discrepancy in evidence from looking and reaching measures. The absence of reaching behaviour in the looking version of the task would reduce the competition between the memory of the hiding location at B trials and the motor traces towards the location A from the previous A trials, which would be present at the reaching version of the task. Moreover, the evidence from studies investigating infants' behaviour in response to a non-hidden object (e.g., object under a transparent cover, Bremner & Knowles, 1984; Butterworth, 1977, or a lids-only condition, Smith et al., 1999) could also be explained under this account. According to this model, the error occurs as a result of an interaction between the motor trace and memory decay among other inputs. The over-activation for a reach towards the location A in the neural activation field would occur when infants experience repetitive reaches towards the location A during A trials. This could be robust enough to override the perceptual input from the non-hidden object. The caveat of this model is its embodied cognition approach which requires no discussions of mental representations (Wilson, 2002). Hence, the role of a hidden object cannot be accounted for under this model, despite the evidence suggesting it is essential for infants' search error (Bremner & Bryant, 2001; Munakata, 1997). Furthermore, the majority of the data used to validate the model were pooled from studies that manipulated the key task elements, such as infants' perceptual or postural input (e.g., Clearfield, Dineva, Smith, Diedrich, & Thelen, 2009; Diedrich, Highlands, Spahr, Thelen, & Smith, 2001; Smith, 2005; Smith et al., 1999) or those that are not published. Therefore, it is difficult to examine how manipulations of these factors might have affected infants' search error in these studies, and to what extent the results from those studies are comparable to results from studies using a standard task procedure originally reported by Piaget (1954). The model validation with the data collected using a standard task procedure would be informative to evaluate to what degree this account can explain the existing reports about the perseverative search error.

Another account that could offer a relatively comprehensive explanation of the reports about the A-not-B error is related to infants' misinterpretation and over-generalisation (Bremner, 1985; Topál et al., 2008). During A trials, infants might generalise the repeated hiding and retrieval of an object from location A and learn that objects are *generally* to be found at location A (Bremner, 1985). This generalisation might be facilitated by a social situation, where the standard A-not-B search task typically occurs. It has been proposed that ostensive cues from their social partner, such as direct eye contact and infant-directed speech, signal infants that the partner is about to transfer relevant and generalisable knowledge to infants (Csibra & Gergely, 2009, 2011; Yoon, Johnson, & Csibra, 2008). Accordingly, ostensive cues induce interpretive bias of generalisability in infants, whereby infants are likely to assume that information communicated with ostensive cues conveys generic information about the referent, rather than episodic information that only concerns under a specific condition (Csibra & Gergely, 2006a, 2006b; Yoon et al., 2008). In the context of the standard A-not-B search task, the adult experimenter typically engages with the infant using various ostensive signals to encourage them to attend to the task. Hence, it is possible that infants misinterpret these social cues from the experimenter, and understand that an object should generally be found at location A (Topál et al., 2008). In order to test the hypothesis, Topál and colleagues (2008) assigned infants into three groups (communicative, non-communicative, and non-social conditions) and assessed the effect of social situations on infants' behaviour at the A-not-B search task. Infants in the communicative conditions engaged in the A-not-B search task with an experimenter who showed positive emotional facial expressions, made a direct eye contact and spoke to the infant in an infant-directed manner. Infants in the non-communicative conditions saw the experimenter sit facing 90 degrees away from the infant and refrain from making an eye contact or verbal interactions with the infant. Infants in the non-social condition saw an object moved, hidden and retrieved through a curtain with no part of the experimenter in their view. In the communicative conditions, 81% of the infants made the A-not-B error, whereas the proportion of infants who made the A-not-B error in the non-communicative and non-social conditions were 48% and 41% respectively. This indicates that social-communicative contexts where ostensive cues are available do influence infants' search performance. Further supporting this social account, Dunn and Bremner (2019) manipulated the experimenters' gaze direction during the delay period to assess how ostensive cues can modulate infants' search performance while avoiding creating an unfamiliar social situation for infants (e.g., an experimenter facing 90 degrees away from the infant when demonstrating the task). Maintaining the ecological validity of the conditions is considered to be important, as such unfamiliarity, as seen in the experimenter's behaviour in Topál and colleagues' (2008) uncommunicative and non-social conditions, could lead to infants' confusion and thus random reaches (Dunn & Bremner, 2019). In this study by Dunn and Bremner (2019), the experimenter first maintained a neutral gaze position during A trials. During the following B trials, they gazed at the correct or incorrect location, or remained neutral in their gaze direction. Infants were more successful when the experimenter looked at the correct location (i.e., B location) on B trials, as compared to when the experimenter gazed neutrally, or at the incorrect location (i.e., A location). The authors further investigated infants' social referencing behaviour after the correct object location was revealed, as an index of infants' expectation regarding the object location (Dunn & Bremner, 2019). It was reported that infants engaged in social referencing more often when they did not find the object at the location they reached for as compared to when they did find the object (Dunn & Bremner, 2019). This suggests that infants do expect to find the object in the location where they searched, as social referencing has been validated as an index of infants' violations of expectation (Dunn & Bremner, 2017). The results provided by Dunn and Bremner (2019) question the view that infants err even when they know where the object actually is (Diamond, 1988). Importantly, the findings demonstrate the social environment plays a critical role in the production of infants' perseverative search errors in the standard version of the search task (Dunn & Bremner, 2019).

Given the empirical support, the social account of search errors is plausible and can explain the role of object permanence, in contrast to the dynamic systems theory. Furthermore, the evidence regarding the error with a non-hidden object (e.g., Bremner & Knowles, 1984; Butterworth, 1977) can also be explained by this account. According to this view, infants are misled by the social situation and misinterpret the experimenter's behaviour during A trials. Therefore, infants believe that some object should be found at location A, regardless of another object being at location B. Nevertheless, receiving helpful social cues does not fully improve infants' search performance in a study by Dunn and Bremner (2019), indicating that infants' misinterpretation in a social situation cannot be a sole cause of the error, and therefore, infants' reaches on the first B trial are still likely to be influenced by other factors, as suggested by the dynamic systems model (Dineva & Schöner, 2018; Smith & Thelen, 1994; Smith et al., 1999; Thelen et al., 2001).

With no existing accounts able to fully explain all the existing reports on the A-not-B errors, further investigation is necessary to fully understand the cognitive processes reflected in this A-not-B error. Taking a multiple-measures approach would be beneficial, given that looking and reaching data have provided opposing reports about where infants might expect objects to be found and decide to reach (Dunn & Bremner, 2019). The use of neural measures is informative to examine infants' cognitive activity (de Haan, 2013; de Haan & Gunnar, 2009). Hence, it could provide another layer of insights into cognitive processes underlying this perseverative error, which are likely to involve information processing, memory, and motor planning. Using a neural measure, it is also possible to examine the online effect of procedural manipulations, which have been commonly done in many of the existing behavioural studies. Task manipulations are essential when using behavioural measures, and carefully planned manipulations can provide invaluable insights into infants' cognition that is reflected in changes in their behaviour. However, the manipulation of task elements can alter infants' cognition at least to some extent. For instance, adding multiple locations can increase memory demand (e.g., Cummings & Bjork, 1981, 1983). Furthermore, introducing an unfamiliar material or extra perceptual or sensory input means that infants need to process additional information as compared to when the traditional procedure is used (e.g., Bremner & Knowles, 1984; Butterworth, 1977). This indicates the possibility that manipulations added to the task procedure could change the level of cognitive load, potentially influencing the whole perceptual, cognitive and memory processes that underlie infant behaviour at the search task. Thus, the results from studies using a modified task procedure have to be carefully examined, in terms of whether the reported “A-not-B error” is comparable to the A-not-B error originally reported by Piaget (1954) and by a number of other researchers. The influence of the procedural manipulation could be tested using a neural measure by directly examining differences in infants’ neural activity across different procedures of the search task.

Despite the potential contribution of neural measures, few neurophysiological studies of infants’ A-not-B errors have been conducted thus far. For instance, Bell and Fox (1992, 1997) examined the relationship between neural activation over frontal areas and infant search performance at the A-not-B search task. They found that the association between successful search and greater EEG power over frontal and occipital regions. Yet, EEG data in these studies were measured while infants quietly sat on their parent’s lap, not whilst they were engaging in the search task. Therefore, the change in EEG power reported in these studies is unlikely to reflect the neural activity *during* the critical period of the task in which information processing, memory and action planning occur. Cuevas and colleagues (2012) investigated infants’ EEG activity during a looking version of the A-not-B search task. Consistent with the reports by Bell and Fox (1992, 1997), they reported an increased EEG coherence (i.e., the synchronised activation between two brain regions) between frontal and other regions during the observation version of the task. However, the task used in this study only required an observation and infants did not need to search for the object by themselves. Therefore, inhibitory control, which has been suggested to be a contributing factor of the error (Diamond, 1985; Smith et al., 1999; Thelen, Schöner, Scheier, & Smith, 2001), was not considered. Furthermore, this study by Cuevas and colleagues (2012) used a multiple-reversal procedure of the A-not-B task, instead of a standard procedure involving one location switch. As mentioned above, this version of the task is likely to be more cognitively demanding in terms of infants’ memory and inhibitory control (Dunn & Bremner, 2019). Therefore, overall, the EEG activation patterns reported in this study might not fully represent the typical neural processes engaged when infants undergo a standard procedure including reach. Using functional near-infrared spectroscopy (fNIRS), Baird and colleagues (2002) measured infants’ brain activity while infants reached for an object. Enhanced activation over frontal regions during the reach was reported, which is consistent with the existing literature using EEG (Bell & Fox, 1992, 1997; Cuevas et al., 2012). However, their task only involved one hiding location, pointing to the limitation of this study that the procedure does not require infants to inhibit a motor trace to reach towards A location based on the memory about the hiding location being B at the B trials. In sum, these studies investigating the neural correlates of infants' successful search performance typically focused on the neural data collected from a situation which lacks in some of the critical components of the A-not-B search task proposed in prior behavioural studies and theories proposed, including an interaction among various inputs, such as hiding location and motor trace, as well as socio-environmental elements (Diamond, 1985; Dineva & Schöner, 2018; Dunn & Bremner, 2019; Smith et al., 1999; Topál et al., 2008).

To conduct research using neural measures that investigates infants' perceptual and cognitive processes underlying infants' behaviour at the standard Piagetian A-not-B search task with no procedure manipulations, we need to establish a novel protocol. To examine any systematic differences in neural activity between the perseverative error and successful search, an EEG oscillatory analysis is useful to explore online cognitive processes (e.g., Hoehl et al., 2014). Oscillatory rhythms observed in EEG signals are thought to be driven by the synchronised activity of large networks of neurons (Jones, Venema, Lowy, Earl, & Webb, 2015). Since the A-not-B error has been proposed to be related to the maturation of frontal cortex (Diamond, 1985, 1988), neural oscillatory power particularly over frontal regions could provide insights into the cognitive processes underlying the A-not-B error.

Among different frequency bands of EEG oscillations, theta- (3-5Hz) and alpha- (6-9Hz) bands are relevant to the cognitive processes proposed to be engaged while infants go through the A-not-B search task, including object knowledge, memory, perceptual and social information processing as well as the planning and execution of reaching behaviour. A study by Begus and colleagues (2015) has shown that theta-band activity can be used as an index of information encoding of novel objects during free exploration (Begus et al., 2015). The authors monitored 11-month-old infants' brain activity whilst they freely explored novel objects. It was reported that enhanced theta-band activity during object exploration was found for objects which were better recognised at the subsequent test phase (Begus et al., 2015). This indicates that theta-band oscillation might reflect the degree of object information encoding and memory. Another study by Begus and colleagues (2016) found that theta-band oscillatory activity can also index infants' expectation formation (Begus et al., 2016). In this study, 11-month-old infants were presented with two types of videos. One of them showed an adult (Informant) who provided information about an object, such as labelling the object and demonstrating its function. The other video showed an adult (Non-informant) who simply pointed at an object or only turned the object around in their hand. Theta-band activity became prominent when the Informant was about to label the object or demonstrate its function, whilst no such effect was observed when the Non-informant was presented. Furthermore, theta-band power was stronger when infants saw a video of adults who spoke their first language than when they saw a video of a foreign language speaker (Begus et al., 2016). Studies with adult population have also reported that theta-band oscillations are prominent while one expects to receive information (Gruber, Watrous, Ekstrom, Ranganath, & Otten, 2013; Guderian et al., 2009). Therefore, theta-band activity in infants might also reflect their anticipation for information (Begus et al., 2016). A similar proposal was made by Orekhova (1999), who found prominent theta-band oscillations while infants engaged in a peek-a-boo game. During peek-a-boo, infants can expect an adult to "reappear" after they "disappear". This also indicates the association between theta-band activity and infants' anticipatory attention (Stroganova, Orekhova, & Posikera, 1999).

Such anticipatory attention in a social situation has also been measured with alpha-band suppression (Orekhova, Stroganova, & Posikera, 2001). Alpha-band activity also holds mu rhythm, typically 6-9Hz, which has been associated with infants' understanding of other people's actions and other social cognitive processes including imitation and theory of mind (Cuevas et al., 2014). As motor representation has been proposed to play a role in the A-not-B error (Landers, 1971; Longo & Bertenthal, 2006), it is possible that alpha-band oscillation, or mu rhythm, might be engaged during the A-not-B error, and perhaps differentiate cognition underlying successful and unsuccessful searches. Further supporting this, whilst the task used was a looking-only and multiple-reverse procedure, Bell (2002) reported that alpha frequency band power successfully differentiated correct and incorrect looking behaviour, in a way that alpha power increased during the correct responses than the incorrect ones.

Based on these, the current study aims to investigate systematic differences in neural activity between infants who make the perseverative error and those who do not, focusing on theta- and alpha-band oscillatory activity. We hypothesise that theta-band power increases when the A-not-B error occurs, reflecting infants' processing of the situation as social, where infants misinterpret an experimenter's intention during the task demonstration and expect to find the object at the location where they reach (Begus et al., 2016; Dunn & Bremner, 2019; Topál et al., 2008). We also expect that alpha-band power difference would index infants' successful and error reach on the B trials, as the cognitive processes underlying the error is likely to include motor planning and inhibition (Bell, 2002; Landers, 1971; Longo & Bertenthal, 2006; Smith & Thelen, 1994; Smith et al., 1999; Thelen et al., 2001).

Methods

Participants

The final sample included in the EEG analysis consisted of 35 infants aged 9 months (14 females, M_{age} = 270.1 days, SD_{age} = 10.2 days). All infants were born full-term (37-41 weeks) and in the normal range for birthweight (over 2500 grams). All infants were recruited through phone calls and emails from a database consisting of parents who expressed an interest in taking part in psychological research. The participating families were predominantly white and middle-class. Additional 15 infants (6 females) were excluded from the final sample due to fussiness (*n*=1), excessive noise in the EEG data (*n*=12), or technical or experimental error (*n*=2). Hence, the data attrition in the current study is 30%. Previous EEG research with infant participants have reported attrition rates between 40 and 65% (e.g. Hoehl, Michel, Reid, Parise, & Striano, 2014; Reid et al., 2004). Hence, this level of exclusion was considered to be typical of this research technique.

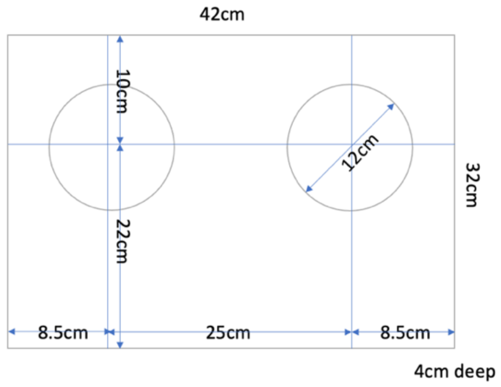
Prior to recruitment, ethical approval was sought and obtained from the Lancaster University Faculty of Science and Technology Research Ethics Committee. All the parents whose child participated in the study were given a full description about the study from the experimenter, and gave a written informed consent on behalf of their child.

Experimental setup and materials

Experimental sessions were conducted in a setup specially designed for the A-not-B search task. A box (42 x 32 x 4cm) with two hiding locations (two 3.5 cm-deep wells, 12cm diameter, 25 cm apart between the centre of the well) was placed on the table (see Figure 6-1). The hiding toy consisted of several plastic loops tied together with one another. Covers to occlude the hiding locations and the toy were made from plain black cloth material (16 x 16cm).

Two camcorders were placed at two different corners of the experimental room, one of which was used to record the infant behaviour and the other the experimenter behaviour.

Figure 6-1. Illustration of the box containing two hiding locations used in the study.



Procedure

Prior to the session, parents were asked to avoid interacting with the infant as much as possible (e.g., to refrain from talking to the child) to ensure parental engagement did not influence infants' behaviour. Infants sat on their parent's lap in front of the experimental setup. Infant EEG was recorded continuously throughout the session.

All infants underwent warm-up trials, the A trials and the B trials. Warm-up trials consisted of three trials, during which the experimenter placed the toy in the central space between the two hiding locations. The object was occluded by a cover each time, and infants were encouraged to reach out for the toy.

Warm-up trials were followed by five A trials. The experimenter lowered the toy in and out of the location A three times, counting audibly to ensure that the infant was attending to the experimenter behaviour as well as the hiding location. Once the toy was placed inside the well, covers were placed over both locations A and B simultaneously. After 5 seconds (a 5-second delay), the hiding stage was pushed towards the infant to allow them to search for the toy. The location of the A trials (left or right) was counterbalanced across participants. The experimenter interacted with the infant in as naturalistic a way as possible. For instance, they maintained eye contact, and verbally encouraged the infant to engage with the search task.

Following the A trials, the B trials commenced. The experimenter lowered the toy in and out of the location B three times, while audibly counting. After that, covers were placed over both locations at the same time to occlude the toy in the hiding location. Following a 5-second delay, the stage was pushed towards the infants to give them an opportunity to search for the toy. The experimenter repeated the B trials until the infant had either completed five trials in total, or were no longer able to sustain their focus on the task. Therefore, the number of B trials varied across participants, with the average being 2.97 B trials (ranging from 1 to 4, $SD_{trials} = 0.97$).

Throughout the session, the infant and the experimenter behaviour were concurrently recorded for offline behavioural coding. The video recordings were also used for the manual rejection of EEG trials in which the infant did not attend to the task, blinked, or moved excessively.

Behavioural analysis

Behavioural coding was conducted modelling Dunn and Bremner (2019) to ensure the current sample engaged in typical A-not-B errors. The behavioural analysis was performed on the data collected from all the infants, regardless of the inclusion/exclusion in the EEG analysis. Two types of behavioural measures were manually coded using the video recordings; (1) the search accuracy on the first B trial as compared to the final A trial ("search accuracy"), and (2) the number of consecutive trials on which infants incorrectly searched during the A and B trials ("error run"). This error run was analysed as an index of the consistency of accurate or inaccurate reaching behaviour (Butterworth, 1977; Dunn & Bremner, 2019).

Search accuracy was analysed using McNemar's Chi^2 test, whilst the error run was assessed using a paired-sample *t*-test or equivalent non-parametric analysis, depending on the normality of the data.

EEG recordings and analysis

Infants' electrical brain activity was recorded using a 128-channel Geodesic Sensor Net. EEG data were amplified with an EGI Net Amps 400 amplifier with a sampling rate of 500Hz (HCGSN 130, EGI, Eugene, OR, USA). Four channels placed on the forehead (channels 125, 126, 127, 128) were disconnected and not used in the current study to avoid excessive discomfort for infant participants, as these electrodes would be positioned near to the eyes. Hence, the data for the current study were collected by 124 channels in total.

EEG was referenced online to the vertex electrode (Cz). In the subsequent offline analyses, an analogue bandpass filter (0.1 to 100 Hz) was applied. Signals (triggers) were sent to the EGI system at the beginning of each 5-second-long delay period ("5-sec delay (period)" hereafter) to segment the data into epochs involving the timeframe of interest (i.e., 5-sec delay).

Initial ERP data preprocessing was performed using NetStation 4.5.4 Waveform Tools. The data were first filtered using 0.3Hz – 30Hz bandpass to remove frequencies not related to the neural activities of interest. The filtered data were then segmented into 5-sec epochs to represent a 5-sec delay period. After the segmentation, the automatic artefact detection tool was applied to each segmented EEG data (i.e., trial). This automatic artefact detection tool marked a channel as bad when the signal amplitude exceeded the set threshold of +/- 200 microvolts for the entire segment, using a moving average of 80ms. The data were also manually and visually inspected, as the algorithm for automatic detection are typically developed for adult EEG data, and not always sensitive to artefact in infant EEG (for further information on infant ERP methods, see Hoehl & Wahl, 2012). Based on the manual artefact detection, trials which included more than 12 channels (i.e., more than 10% of the channels used for data acquisition) containing excessive noises were excluded for the following steps of the analysis.

Following the artefact detection, the EEG data were divided into four conditions according to infants' search performance (accurate or inaccurate search) on each trial (A or B trials). As expected from the literature, infants in the current sample were more successful in retrieving the toys at A trials than B trials. Therefore, more trials were found for the A trial accurate search condition (n = 57, "A accurate" hereafter), than other conditions (A trial inaccurate search, or "A inaccurate", n = 11; B trial accurate search, or "B accurate", n = 19; and B trial inaccurate search, or "B inaccurate", n = 23). In power analysis, however, a condition with a smaller number of trials is more likely to show larger effects (Cohen, 2014). Hence, it was decided to only include the fourth and fifth A trials as the A accurate condition, the rationale being that the definition of the A-not-B error specifies the accurate search at the last (i.e., 5th) A trial and the inaccurate search at the first B trial. As a result, the number of infants who contributed final data to each condition was *n* = 22 (32 trials in total) for A accurate, *n* = 14 (23 trials in total) for A inaccurate, *n* = 10 (27 trials in total) for B accurate, and *n* = 14 (25 trials in total) for B inaccurate (see Table 6-1). Not all the infants contributed to all the conditions, as some infants only contributed to the A accurate but not to the other three conditions. Whilst it is common to only include infants who contributed data to all the conditions in the final analysis in infant EEG research, due to the nature of the study design, the modification of inclusion criteria was necessary in this study. In conventional EEG studies, the same stimulus is typically presented multiple times to increase the number of noise-free trials for each condition per infant (de Haan, 2013).

In the current study, however, each infant only engaged in five A trials and up to four B trials in a live interactional setting. This resulted in high attrition of the data, and we could only include one or a few usable trials per condition per infant. Nevertheless, prior studies have shown that a reliable EEG analysis can be done with as few as three trials per condition per infant (Kaduk et al., 2013; Stets & Reid, 2011). As the following analysis was mainly to show the feasibility of this study procedure, it is considered to be acceptable to proceed with one or two trials per infant.

Table 6-1. The number of participants and trials included in the final analyses for each condition. The final sample for the A accurate condition only consisted of the fourth and fifth A trials.

	A accurate	A inaccurate	B accurate	B inaccurate
Trials included	32	23	27	25
Participants included	22	10	23	14

Artefact-free segments were then submitted to fast Fourier transformations (FFTs). The EEG power was computed over the 5-second delay (from time 0 to 5,000ms relative to the onset of the 5-sec delay period), using a Hanning-tapered window of 555ms with 50% overlap. Power estimates were extracted between 0 and 30Hz in steps of 1Hz. Grand averages of the FFTs were computed for each individual. For the statistical tests, the $\log_{10}$ transformation was performed over the data to improve the normality of the data distribution.

Frequency ranges were specified to enable a direct comparison between the current study and prior research (Begus et al., 2015; St. John et al., 2016). To assess differences across conditions, the percentage of the amplitude within the frequency band of interest (e.g., theta, 3-5Hz) relative to the total power (0.1-30Hz) was examined. The relative measure was considered to be more robust towards between-subject variability in terms of the overall amplitude, and also optimal when spectral values are not comparable across different frequency bands (Benninger, Matthis, & Scheffner, 1984; Clarke, Barry, McCarthy, & Selikowitz, 2001; Orekhova, Stroganova, & Posikera, 2001).

Based on the existing literature, the fronto-central electrode sites were analysed to investigate differences in neural activation within theta frequency band (3-5Hz) (Begus, Southgate, & Gliga, 2015; Orekhova, Stroganova, Posikera, & Elam, 2006; Saby & Marshall, 2012). Other frequency ranges (4-6Hz and 6-9Hz) over frontal, central, temporal and parietal regions were also analysed based on prior research investigating the neural representation of the processing of social stimuli (Jones et al., 2015; St. John et al., 2016). The electrodes included in each region of interest are summarised in Table 6-2.

Table 6-2. The electrodes of a 128-channel Geodesic Sensor Net used for each analysis.

Channel regions	Channel numbers (128-channel Geodesic Sensor Net)
Fronto-central regions (3-5Hz)	4 9 10 11 15 16 18 19 22
Frontal regions (4-6Hz, 6-9Hz)	3 4 5 10 12 18 19 20 23 24 27 28 33 34 38 116 117 118 121 122 123 124
Central regions (4-6Hz, 6-9Hz)	29 30 26 41 87 103 104 105 111
Temporal regions (4-6Hz, 6-9Hz)	39 43 44 45 48 49 50 56 57 58 63 64 95 96 99 100 101 107 108 113 114 115 119 120
Parietal regions (4-6Hz, 6-9Hz)	42 46 47 51 52 53 54 59 60 61 67 77 78 79 85 86 91 92 93 97 98 102

Results

Behavioural analyses on infants' perseverative reaches

Search accuracy

The sample for the analysis of search accuracy consisted of 27 infants. Infants who reached to both locations, or neither location on at least either the final A trial or the first B trial were excluded from the analysis (*n*=8).

Table 6-3 shows the result of McNemar's *Chi*^{2} test, which indicated no statistically significant difference in the proportion of infants who made the A-not-B error as compared to those who did not err (*p* = .238). This implied that the current data might not reflect the standard occurrence of infants' A-not-B search error.

Table 6-3. Search accuracy on the final A trial and the first B trial.

	Final A trial	
First B trial	Accurate	Inaccurate
Accurate	10	5
Inaccurate	11	1

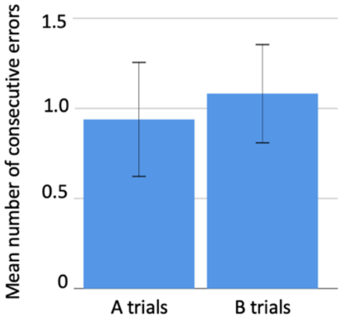
Note. The bottom left corner (Final A trial accurate and first B trial inaccurate) represents the number of infants who made the A-not-B error.

Error run

The sample for the analysis of error run consisted of 39 infants. This error run analysis included infants who made a 'random reach' at either the final A or the first B trial, of whom data were excluded in the previous analysis on the search accuracy. The 'random reach' was defined as an infant's reaching to both or neither of the two locations. These were not considered as 'inaccurate' searches, as such reaching behaviour is better interpreted as incidental, rather than an infant's intentional choice. Therefore, the number of these random searchers were not counted as 'accurate' or 'inaccurate' in the error run count. This was considered appropriate for the focus of this analysis being the number of consecutive errors intentionally made.

Figure 6-2 illustrates the mean number of consecutive errors made during the A trials as well as the B trials, with the error bars representing the standard error. A Shapiro-Wilk test indicated that neither of the A- or B-trials data met the assumption of normality ($p < .001$, $p < .001$ respectively). Therefore, a non-parametric test was used to evaluate a difference across trials. A Wilcoxon signed-rank test found no significant difference between the error runs during the A trials ($M_{error\text{-}run} = 0.38$, $SD_{error\text{-}run} = 0.62$) and B trials ($M_{error\text{-}run} = 1.11$, $SD_{error\text{-}run} = 0.95$) ($Z = -1.44$, $p = .15$). Consistent with the search accuracy measure, this error run analysis also indicated that our current data may not reflect the standard occurrence of the A-not-B error that had been reported in prior studies.

Figure 6-2. The mean length of error made in a row during A and B trials.



Note. The error bars represent the standard error.

Comparison with a prior study

To further examine whether the data in the current study represented the typical A-not-B error, the results of the current behavioural analyses were compared to that of Dunn and Bremner (2019), for the similarity in terms of a setup, task procedure, and behavioural measures used. Compared with their results, as shown in Tables 6-4 and 6-5, infants in the current study appeared less likely to reach the accurate location on the last A trial, and more likely to reach the accurate location on the first B trial. This suggests that infants' reaching behaviour observed in the current sample might have occurred rather incidentally, and the A-not-B error at random.

Table 6-4. Comparison of McNemer's Chi^2 score test result with a prior study by Dunn and Bremner (2019)

	Dunn & Bremner (2019)	Current study
A5 accurate - B1 accurate	3	10
A5 accurate - B1 inaccurate	12	11
A5 inaccurate - B1 accurate	0	5
A5 inaccurate - B1 inaccurate	1	1

Table 6-5. Comparison of the error runs during the A and B trials between Dunn & Bremner (2019) and the current study.

	Dunn & Bremner (2019)		Current study	
	A trials	B trials	A trials	B trials
M_{trials}	0.38	2.74	0.78	1.11
SD_{trials}	0.62	1.44	0.90	0.95
Error occurrence (%)	7.60	71.91	15.6	37.37

Note. The average number of B trials was 3.81 in Dunn and Bremner (2019), and 2.97 in the current study. Both studied conducted five A trials for each infant. Error in percentage was calculated by dividing the mean error run by the mean total number of B trials performed. However, this is for comparative purposes only and does not reflect the proportion of error occurring during the A or B trials.

Summary of the behavioural analyses

The results of the behavioural analyses indicated that the data in the current study were unlikely to be a typical representation of the A-not-B search error, in terms of the change in search accuracy from the last A trial to the first B trial, and also the number of continuous errors made during each set of A and B trials. More specifically, infants who participated in the current study appeared to have made more errors during the A trials, and fewer errors during the B trials.

Despite the behavioural results suggesting the lack of typical A-not-B errors, there were accurate and inaccurate search behaviours made during the A and B trials, which provided us with a sufficient number of analysable (i.e., artefact-free) trials for EEG analyses to explore neural activation patterns underlying infants' search performance. Therefore, the EEG analyses using the current data were considered to be practicable, which will be reported in the next group of sections.

EEG analyses on neural activity during a delay period before the search behaviour

Based on the prior literature, our examination focused on theta- and alpha- frequency ranges; namely, (1) theta-band (3-5Hz) activity over fronto-central electrode regions, which was previously shown to mark anticipatory learning (Orekhova, Stroganova, & Posikera, 1999), and the better encoding and recognition of objects (Begus et al., 2015), (2) theta-band (4-6Hz) activity over frontal, central, temporal and parietal regions, which has been associated with neural activities underlying social interactions (Jones et al., 2015; St. John et al., 2016), as well as (3) alpha-band (6-9Hz) activity over frontal, central, temporal and parietal regions, which has been linked to social information processing (Jones et al., 2015; St. John et al., 2016).

Figure 6-3 compares the relative EEG power (i.e., the power within the frequency range of interest relative to the total power) in the frequency bands of interest across conditions. Visual examination of the data (Figure 6-3) indicated that there seemed to be no evident differences across conditions in terms of theta oscillatory activity over central electrode sites, as well as alpha frequency range (6-9Hz) over any regions examined. Therefore, the following sections will focus on reporting the results of frequency analyses of theta-band activity over fronto-central, frontal, temporal and parietal electrode regions. Descriptive statistics of the data analysed were summarised in Table 6-6, and Figure 6-4 depicts bar charts comparing the power within the theta-band frequency range across conditions, over fronto-central, frontal, temporal and parietal regions.

Figure 6-3. Comparisons across the four conditions (A accurate [dark blue], A inaccurate [light blue], B accurate [green], and B inaccurate [yellow]), in terms of the percentage of the power in the frequency ranges of focus relative to the total frequency range (0.1-30Hz).

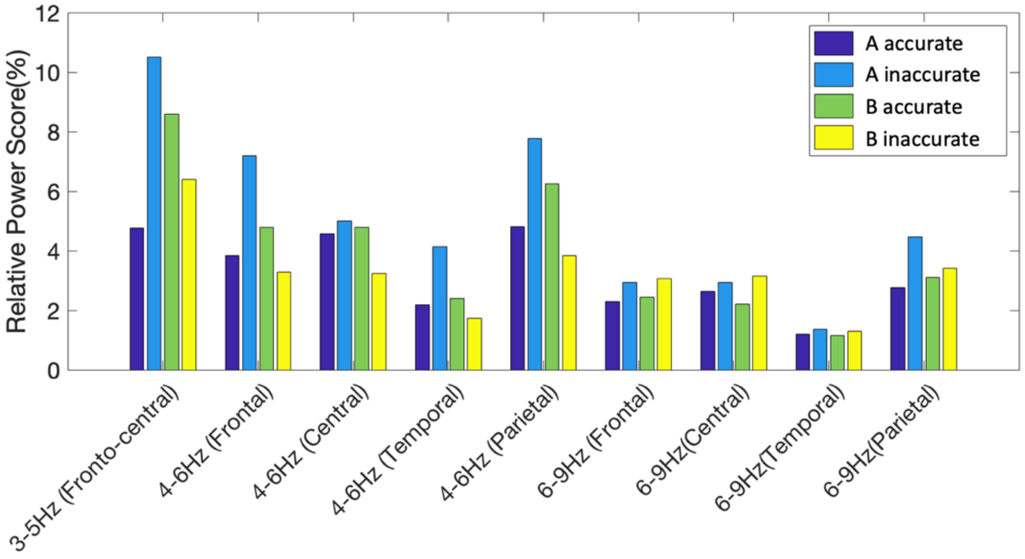
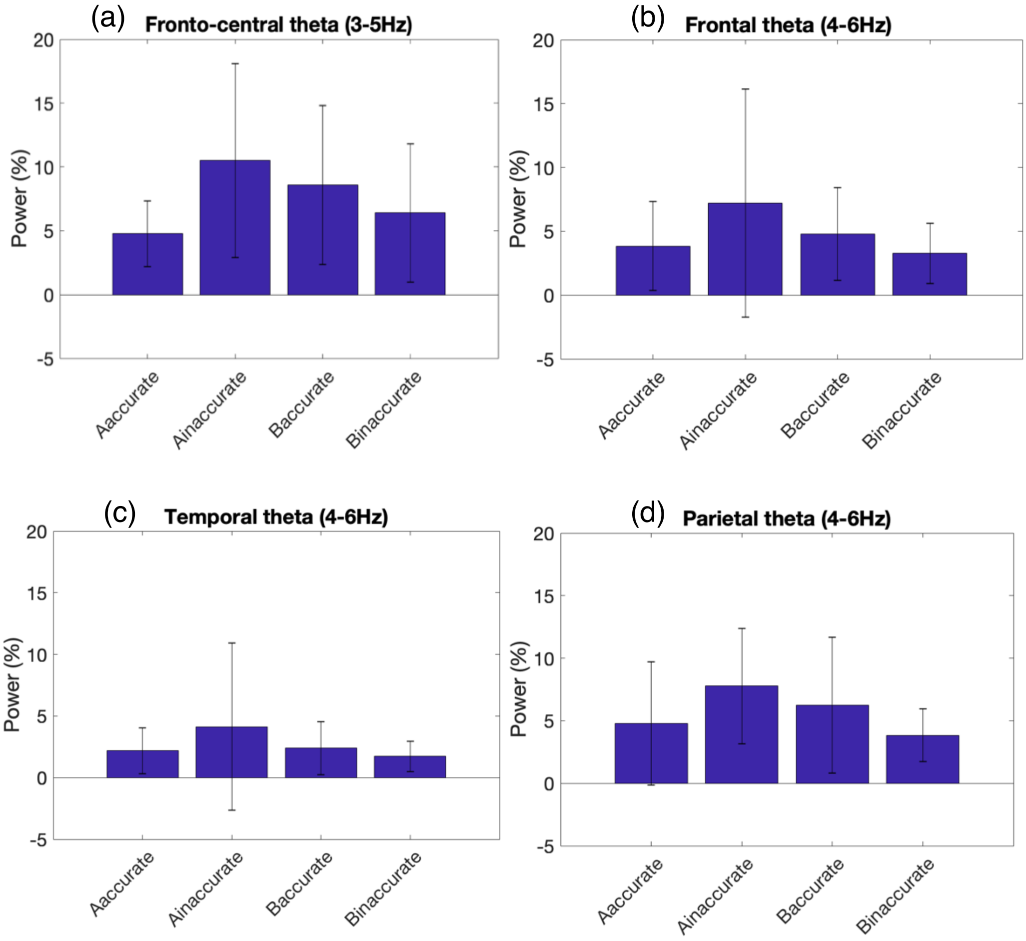


Table 6-6. Descriptive statistics of the data for theta-band oscillatory activity over fronto-central, frontal, temporal and parietal regions.

		N	*M*	*SD*
Fronto-central	A accurate	22	4.77	2.55
	A inaccurate	10	10.51	7.61
	B accurate	23	8.60	6.21
	B inaccurate	14	6.40	5.39
Frontal	A accurate	22	3.85	3.50
	A inaccurate	10	7.20	8.94
	B accurate	23	4.79	3.62
	B inaccurate	14	3.27	2.35
Temporal	A accurate	22	2.18	1.84
	A inaccurate	10	4.14	6.78
	B accurate	23	2.39	2.16
	B inaccurate	14	1.72	1.22
Parietal	A accurate	22	4.81	4.92
	A inaccurate	10	7.78	4.61
	B accurate	23	6.25	5.42
	B inaccurate	14	3.85	2.12

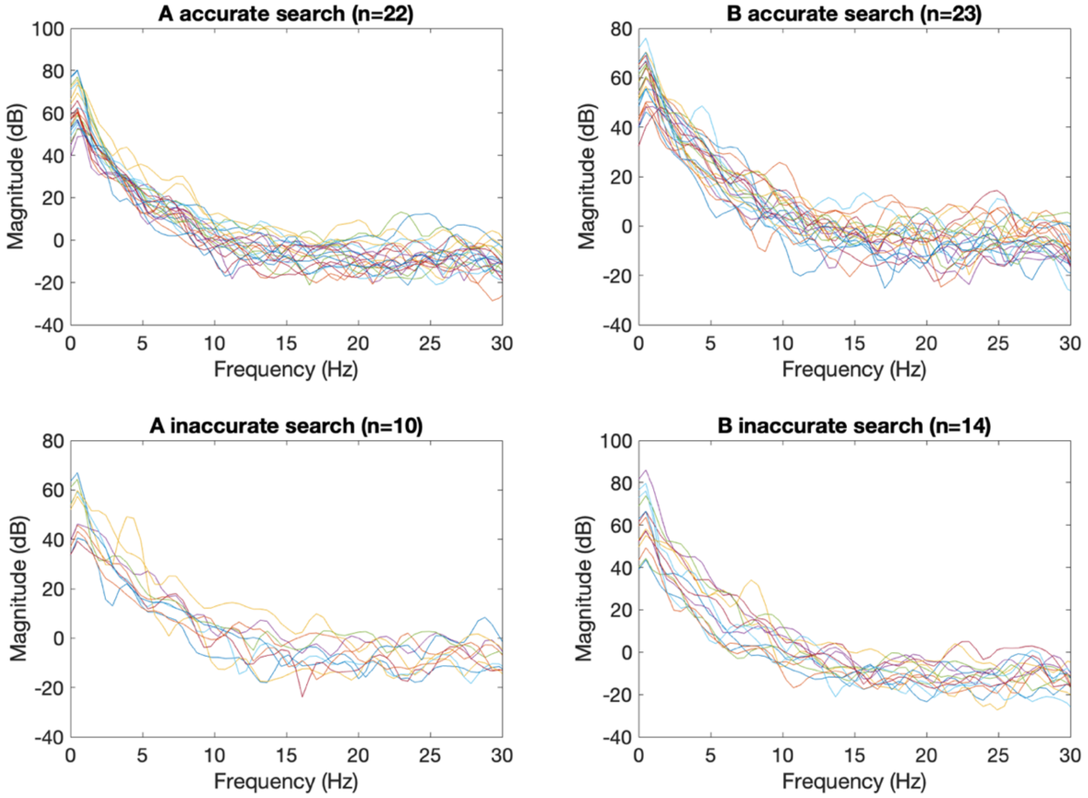
Figure 6-4. Bar charts comparing the power of theta-band oscillations across conditions, over (a) fronto-central, (b) frontal, (c) temporal and (d) parietal regions.



Note. The frequency band analysed was 3-5 Hz for the fronto-central regions, and 4-6Hz for other regions, to enable for a direct comparison with prior studies (Begus et al., 2015; St John et al., 2016). Error bars represent standard deviation.

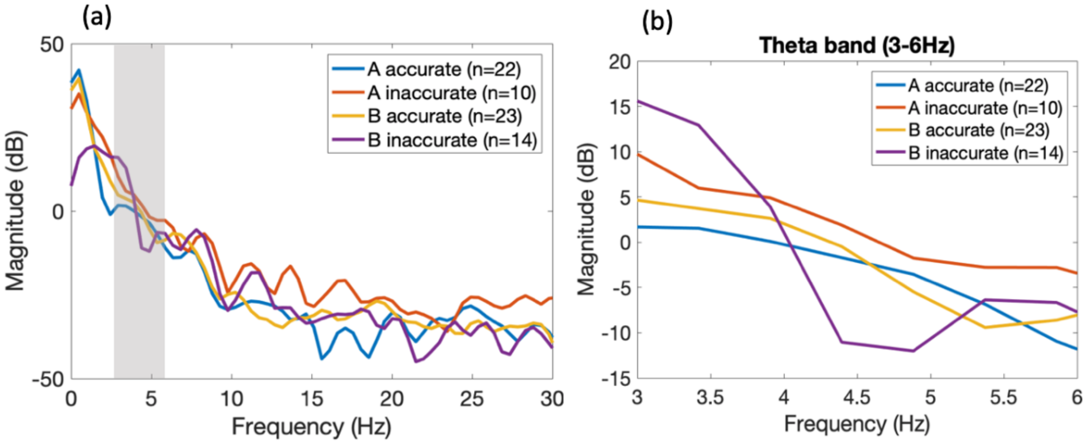
Figure 6-5 shows the individual power spectra (i.e., the average over the usable trials of each participant) across conditions, computed by fast Fourier transformations. Figures 6-6 represents the average power spectra across participants across conditions. A single line in Figure 6-6 represents a group average over individual lines shown in Figure 6-5, in which each line represents an individual participant average.

Figure 6-5. Illustrative examples of the results of frequency analyses.



Note. Each line represents a power spectrum of the averaged power over the fronto-central electrode sites of each participant, for the A accurate condition (top left), the A inaccurate condition (bottom left), the B accurate condition (top right) and the B inaccurate condition (bottom right).

Figure 6-6. Illustrations of average power spectra across participants for the A accurate (blue), the A inaccurate (orange), the B accurate (yellow) and the B inaccurate (purple) conditions.



Note. The left figure (a) shows the power spectra across the entire frequency range of the data in the current study, while the right figure (b) only depicts the power spectra within the theta frequency range (3-6Hz). The grey shading in the left graph (a) represents a theta frequency range (3-6Hz), which is displayed in the right graph (b).

After the exclusion of EEG trials with excessive noise, no infants contributed data to all the four conditions, and the sample sizes across conditions were not equal to one another. Therefore, planned ANOVAs were considered to be inappropriate. Instead, we conducted three separate statistical comparisons between two groups of interest: namely, (1) A trial accurate and A trial inaccurate reaches (*n* = 3), (2) B trial accurate and B trial inaccurate reaches (*n* =10), and (3) A trial accurate and B trial accurate reaches (*n* =15). Whilst a comparison between A inaccurate and B inaccurate trials was planned, there was only one infant who contributed data to both conditions. Hence, we were unable to conduct a reliable statistical assessment on differences in theta frequency activity between A inaccurate and B inaccurate trials. Despite multiple comparisons performed, we did not correct the error rate using a conventional method (e.g., Bonferroni correction). The current data were collected as a pilot study for a future research project investigating the neural mechanisms underlying the A-not-B search performance. Not correcting the error rate was considered optimal to maximise freedom to formulate hypotheses for future investigation (Goeman & Solari, 2011; Saville, 1990). For each of the three comparisons, the result of the analysis on the fronto-central theta (3-5Hz) oscillatory activity is reported first. This is followed by the assessment of the frontal theta (4-6Hz) oscillations, after which the temporal theta (4-6Hz) oscillations are examined. Lastly, the result of the analysis on the parietal theta (4-6Hz) oscillation is described.

The results of a normality test (Shapiro-Wilk test) for each comparison were summarised in Table 6-7. As none of the pairs met the assumption of normality, a non-parametric comparison (a Wilcoxon singed-rank test) was used as a statistical assessment across conditions.

Table 6-7. Results of the normality tests (Shapiro-Wilk test).

A accurate and A inaccurate		*n*	Statistic	*df*	*p*-value
Fronto-central	A accurate	3	.94	3	.511
	A inaccurate	3	.79	3	.081
Frontal	A accurate	3	.10	3	.894
	A inaccurate	3	.10	3	.883
Temporal	A accurate	3	.10	3	.878
	A inaccurate	3	.89	3	.355
Parietal	A accurate	3	.77	3	.040
	A inaccurate	3	.92	3	.463

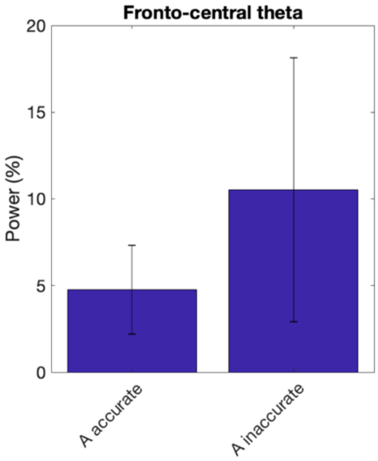
B accurate and B inaccurate		*n*	Statistic	*df*	*p*-value
Fronto-central	B accurate	10	.91	10	.289
	B inaccurate	10	.81	10	.020
Frontal	B accurate	10	.91	10	.295
	B inaccurate	10	.90	10	.222
Temporal	B accurate	10	.77	10	.006
	B inaccurate	10	.75	10	.003
Parietal	B accurate	10	.86	10	.076
	B inaccurate	10	.94	10	.597

A accurate and B accurate		*n*	Statistic	*df*	*p*-value
Fronto-central	A accurate	16	.95	16	.511
	B accurate	16	.85	16	.012
Frontal	A accurate	16	.70	16	<.001
	B accurate	16	.89	16	.054
Temporal	A accurate	16	.84	16	.008
	B accurate	16	.65	16	<.001
Parietal	A accurate	16	.95	16	.430
	B accurate	16	.76	16	.001

Accurate and inaccurate trials during A trials

Fronto-central theta (3-5Hz) oscillatory activity. A Wilcoxon signed-rank test (*n*=3) indicated no significant difference across conditions (*p*=.109).

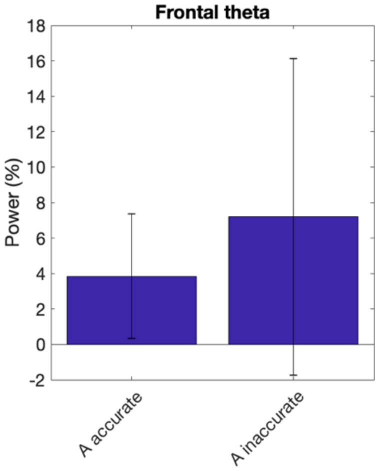
Figure 6-7. Bar chart comparing the power of theta-band activity across conditions over fronto-central electrode sites.



Note. The frequency extracted was a range of 3 to 5Hz. Error bars represent the standard deviation. The data represented in the bar chart include all the data available, and are not identical to the data used in the statistical assessment.

Frontal theta (4-6Hz) oscillatory activity. A Wilcoxon signed-rank test (*n*=3) suggested that there was no difference across conditions (*p*=.109).

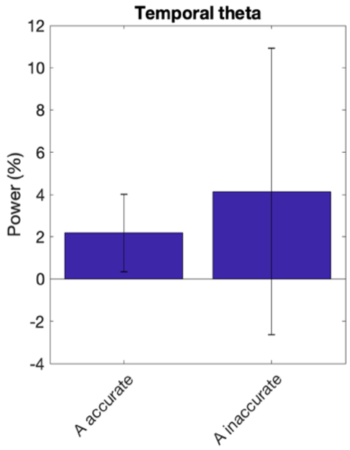
Figure 6-8. Bar chart comparing the power of theta-band activity across conditions over frontal electrode sites.



Note. The frequency extracted was a range of 4 to 6Hz. Error bars represent the standard deviation. The data represented in the bar chart include all the data available, and are not identical to the data used in the statistical assessment.

Temporal theta (4-6Hz) oscillatory activity. A Wilcoxon signed-rank test (*n*=3) found no significant difference across conditions (*p*=.593).

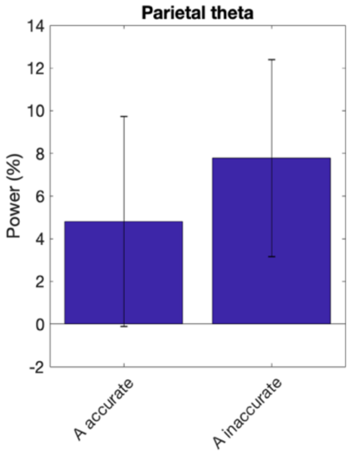
Figure 6-9. Bar chart comparing the power of theta-band activity across conditions over temporal electrode sites.



Note. The frequency extracted was a range of 4 to 6Hz. Error bars represent the standard deviation. The data represented in the bar chart include all the data available, and are not identical to the data used in the statistical assessment.

Parietal theta (4-6Hz) oscillatory activity. A Wilcoxon signed-rank test (*n*=3) indicated no significant difference across conditions (*p*=.109).

Figure 6-10. Bar chart comparing the power of theta-band activity across conditions over parietal electrode sites.

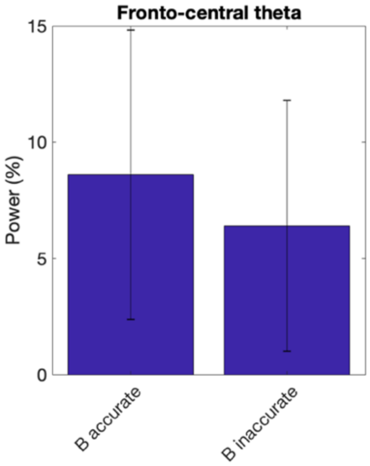


Note. The frequency extracted was a range of 4 to 6Hz. Error bars represent the standard deviation. The data represented in the bar chart include all the data available, and are not identical to the data used in the statistical assessment.

Accurate and inaccurate search during B trials

Fronto-central theta (3-5Hz) oscillatory activity. A Wilcoxon signed-rank test (*n*=10) indicated no significant difference across conditions (*p* =.878).

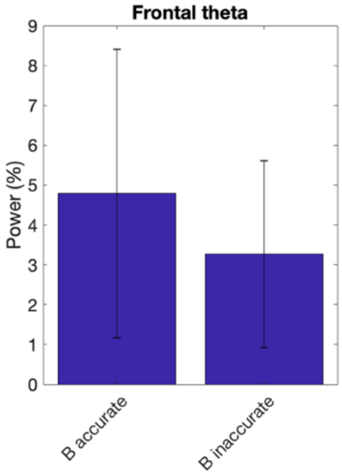
Figure 6-11. Bar chart comparing between theta oscillatory activity over fronto-central channel regions occurring before B accurate (left) and inaccurate search (right).



Note. Error bars represent the standard deviation. The data represented in the bar chart include all the data available, and are not identical to the data used in the statistical assessment.

Frontal theta (4-6Hz) oscillatory activity. A Wilcoxon signed-rank test (*n*=10) found no significant difference across conditions (*p* =.646).

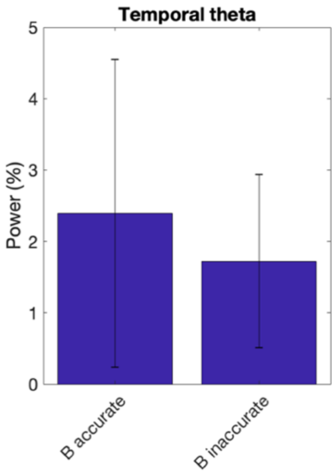
Figure 6-12. Bar chart comparing between theta oscillatory activity over frontal channel regions occurring before B accurate (left) and inaccurate search (right).



Note. Error bars represent the standard deviation. The data represented in the bar chart include all the data available, and are not identical to the data used in the statistical assessment.

Temporal theta (4-6Hz) oscillatory activity. A Wilcoxon signed-rank test (*n*=10) suggested no significant difference across conditions (*p* =1.00).

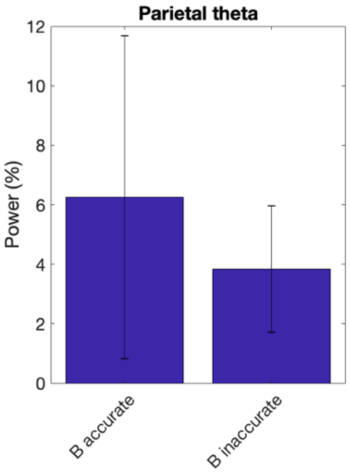
Figure 6-13. Bar chart comparing between theta oscillatory activity over temporal channel regions occurring before B accurate (left) and inaccurate search (right).



Note. Error bars represent the standard deviation. The data represented in the bar chart include all the data available, and are not identical to the data used in the statistical assessment.

Parietal theta (4-6Hz) oscillatory activity. A Wilcoxon signed-rank test (*n*=10) indicated no significant difference across conditions (*p* =.33).

Figure 6-14. Bar chart comparing between theta oscillatory activity over parietal channel regions occurring before B accurate (left) and inaccurate search (right).

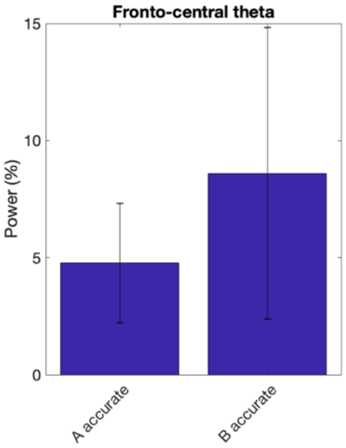


Note. Error bars represent the standard deviation. The data represented in the bar chart include all the data available, and are not identical to the data used in the statistical assessment.

Accurate search during A and B trials

Fronto-central theta (3-5Hz) oscillatory activity. A Wilcoxon signed-rank test (*n*=16) found no significant difference across conditions (*p* = .121).

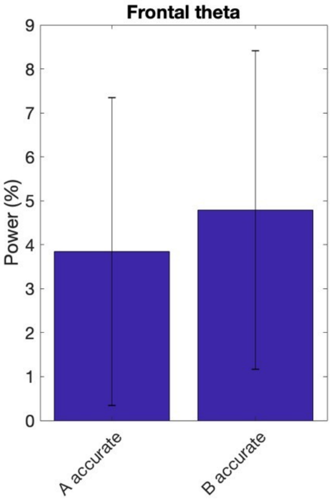
Figure 6-15. Bar chart comparing between theta oscillatory activity over fronto-central channel regions occurring before A accurate (left) and B accurate search (right).



Note. Error bars represent the standard deviation. The data represented in the bar chart include all the data available, and are not identical to the data used in the statistical assessment.

Frontal theta (4-6Hz) oscillatory activity. A Wilcoxon signed-rank test (*n* =16) indicated no significant difference across conditions (*p* = .124).

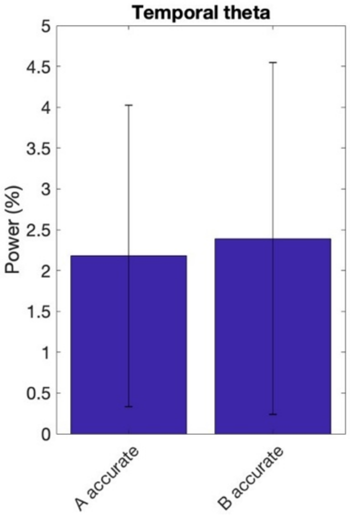
Figure 6-16. Bar chart comparing between theta oscillatory activity over frontal channel regions occurring before A accurate (left) and B accurate search (right).



Note. Error bars represent the standard deviation. The data represented in the bar chart include all the data available, and are not identical to the data used in the statistical assessment.

Temporal theta (4-6Hz) oscillatory activity. A Wilcoxon signed-rank test (*n*=10) indicated no significant difference across conditions (*p*= .816).

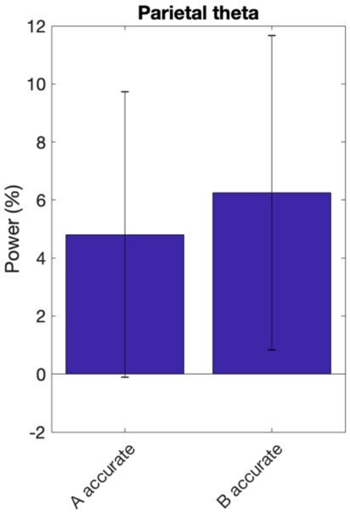
Figure 6-17. Bar chart comparing between theta oscillatory activity over temporal channel regions occurring before A accurate (left) and B accurate search (right).



Note. Error bars represent the standard deviation. The data represented in the bar chart include all the data available, and are not identical to the data used in the statistical assessment.

Parietal theta (4-6Hz) oscillatory activity. A Wilcoxon signed-rank test (*n* =10) found no significant difference across conditions (*p* = .469).

Figure 6-18. Bar chart comparing between theta oscillatory activity over parietal channel regions occurring before A accurate (left) and B accurate search (right).



Note. Error bars represent the standard deviation. The data represented in the bar chart include all the data available, and are not identical to the data used in the statistical assessment.

Discussion

Given that the existing evidence suggests a multitude of factors that influence infant behaviour on the Piagetian A-not-B search task, and that different measures could highlight different cognitive processes, it is considered optimal to utilise multiple measures when examining why infants make the perseverative errors (Dunn & Bremner, 2019). To establish a measure that can directly explore infants' cognition that emerges during a critical time of the task, this study utilised an EEG as a measure and investigated infants' neural activity during the A-not-B search task, using a traditional procedure. The use of neural measures could also enable the comparison between different procedures in terms of the cognitive processing, and potentially help us explain the disparity in the existing reports. We collected and analysed 9-month-old infants' EEG oscillatory activity in theta- and alpha-bands during the delay period. However, the behavioural analysis indicated that the current sample of infant participants did not show a standard perseverative search error pattern. This has two implications; that the neural data obtained might not reflect accurate cognitive processes during the typical A-not-B search task, and that there were fewer data points than expected which reflect the typical infant behaviour at this task, across conditions and accuracy of reaches. EEG data were nonetheless analysed to look for any potential differences between accurate and inaccurate searches overall in terms of the neural activity engaged during the delay period of the task. Whilst data indicated that theta-band oscillatory activity could be a potential neural marker of infant search performance, the large individual differences observed in the current sample likely contributed to a lack of significant effect.

Behavioural analyses indicated that our current data did not reflect the typical A-not-B search error behaviour in terms of the frequency of each event (i.e., accurate or inaccurate search at A and B trials). A comparison with a prior study suggested that infants in the current study made more errors during the A trials, and fewer errors during the B trials. Further inspection of video recordings indicated a likely cause in a potential issue with the apparatus. For many infants, the height of the box used to hide an object may have been very close to their eye level, and thus it could have been difficult for them to see inside the hiding wells. Hence, these infants could not see the object (and the end goal) when it was placed in the well. This could lead to confusion about the task, leading to random reaching. An analogous example could be drawn from a study by Boyer and colleagues (2011), where they used mechanical claws to demonstrate the A-not-B search task. This study found that only when infants were shown the function of mechanical claws did they commit the A-not-B error. Infants who were familiarised with mechanical claws but were not shown what the function was did not make the A-not-B error (Boyer et al., 2011). This result indicates that infants' understanding of what an experimenter is trying to do is a core element of the A-not-B error, consistent with the dynamic systems view (Dineva & Schöner, 2018; Smith & Thelen, 1994; Smith et al., 1999; Thelen et al., 2001) as well as the social account of this search error (Dunn & Bremner, 2019; Topál et al., 2008).

Despite the behavioural analyses suggesting the lack of typical A-not-B search error patterns in the current sample, there were accurate and inaccurate searches during the A and B trials. The number of these searches was sufficient to perform preliminary analysis of neural activation patterns underlying accurate and inaccurate reaches overall. By analysing the data, we aimed to offer a proof of concept for the analysis of EEG activity during a standard A-not-B search task paradigm, comparing across four different search conditions: accurate A trials (A accurate), inaccurate A trials (A inaccurate), accurate B trials (B accurate), and inaccurate B trials (B inaccurate).

We found no evidence for a significant difference across conditions, in term of theta- or alpha-band oscillatory power over the fronto-central (theta only), frontal, parietal, temporal or occipital electrode regions. This null result may be due to a large standard deviation in data for each condition. This is likely to reflect considerable individual differences in the current data, in terms of neural activation in theta and alpha frequency bands. This large standard deviation is likely to be related to our small sample size and the small number of EEG trials included in the EEG analysis. Typically, studies using EEG average the data over a number of trials in the same condition to improve the signal to noise ratio (Luck, 2005). In infant studies, the number of trials averaged can be three but generally over ten (Stets & Reid, 2011). To keep this standard and enable data averaging, the current study grouped the obtained data into four conditions, according to the trial (A or B) and the search performance (accurate or inaccurate). Due to the nature of the task, however, data from each search could be different to some extent even within the same condition in terms of infants' cognition involved. Specifically, it has been proposed that the motor trace strengthens as infants repeat the same reaching behaviour during A trials (Munakata, 1998; Thelen et al., 2001). Furthermore, it has been reported that infants learn to reach correctly as they repeat B trials (Cummings & Bjork, 1981, 1983). Hence, very strictly speaking, the data averaged together in the current study may naturally contain some variation. Moreover, the current data were collected while infants engaged in a naturalistic, live, and dynamic interaction. In a typical experimental paradigm, fluctuating variance could be a function of sample sizes that are too small or other factors that are unintentionally confounding investigation. Therefore, it is often effective to ensure experimental control is rigid and strict to reduce the variance. Yet, such control did not accord with the focus of this study which aimed to keep the social interactional format of the traditional A-not-B search task procedure. This attempt might have contributed to a large standard variation in the current data.

An attempt to collect EEG data during a live interactional setting was feasible as the current study has shown, yet challenges remain. This has resulted in a large attrition rate for the data. This made a planned analysis using a repeated measures ANOVA unsuitable due to the unequal sample size for each condition. Hence, we performed exploratory comparisons including between B accurate and B inaccurate, and between A accurate and B accurate trials. The comparison between B accurate and B inaccurate searches was critical for our ultimate research aim of investigating systematic differences between infants who err and those who do not at B trials. If there should be any difference in neural activity occurring during a 5-second delay period before accurate and inaccurate searches during the B trials, that could potentially indicate cognitive processes that differentiate infants' successful and unsuccessful reaching on B trials. The comparison of neural activations occurring prior to accurate searches between the A and B trials could potentially identify cognitive differences that could explain why infants err on B trials but not on A trials. Nonetheless, the current data found no significant difference across conditions, likely due to the large variance in the data, as discussed above. Whilst we also compared A accurate trials with A inaccurate trials, this comparison would not have been done if the participants had shown a typical AB search pattern and made few errors during the A trials.

Overall, the behavioural result of the current sample was a critical limitation of the current study in many respects. This means that the neural results reported here might differ from what could have been obtained based on neural data collected in a more typical A-not-B search behavioural paradigm. Nevertheless, this study has provided evidence that using EEG as a measure of cognitive activity underlying the Piagetian A-not-B search error is feasible without interfering with the standard task procedure. This means that it is also possible to use this data collection and analysis pipeline for a modified version of the A-not-B task procedure, enabling us to compare how different task manipulation can influence infants' cognition. We believe this study has provided a platform for future research which aims to untangle contradicting evidence reported thus far. We hope that this will inspire further investigation of the perseverative error, and discussion on what this error tells us about infant cognition and its development.

Acknowledgement

The study reported here is a secondary data analysis. I thank Dr Kirsty Dunn for giving me the permission to analyse the data she previously collected with her colleagues. The study was designed by Dr Kirsty Dunn and Professor Gavin Bremner. The data were acquired by Dr Szilvia Linnert with the assistance of Martina Arioli. The data had not been analysed prior to the current study. I extend my thanks to Barrie Usherwood for his support with MATLAB coding.

Author Contribution

Sayaka Kidby: analysis design, analysis, writing, review. Vincent Reid: analysis design, review. Szilvia Linnert: data acquisition. Gavin Bremner: study design. Kirsty Dunn: study design, analysis design, review.

Chapter 7

General discussion

Revisiting the theoretical basis of the thesis objectives

Infants are not passive receivers of information; they participate in social interactions (Cohn & Tronick, 1988; Fogel et al., 1982; Tronick et al., 1980). The existing evidence consistently suggests that infant behaviour, such as looking and pointing, is both influenced by and influences other people's behaviour in a social situation, creating mutual and reciprocal regulation systems (Field, 1978; Tronick, 1989). What neural mechanisms underlie such infants' active engagement in social interactions is one of the core questions that need addressing so we can better understand how infants develop their socio-cognitive skills which enable them to competently engage with others. Yet, the conventional approach of infant social cognition research utilising neural measures typically involves the use of a paradigm where participants only act as observers of social stimuli, taking a 'third-person' perspective, 'isolated' from social contexts (Becchio et al., 2010). In such a paradigm, infants are treated as passive observers, rather than active participants of reciprocal social interactions. The use of such paradigms where little or no social elements are present is largely due to the technological constraints of neural measures, such as vulnerability to motion. However, this means we have not yet fully addressed what neural processes underlie infants' active engagement in reciprocal social interactions, and how the underlying social-cognitive skills develop.

With these questions unanswered, this thesis has argued for a transition of infant neurophysiological research from a controlled 'third-person' paradigm to a more dynamic and interactional 'second-person' paradigm, where participants can actively engage in a social interaction with others rather than simply observe other people from outside of the social situation (Schilbach et al., 2013). The importance of moving towards the second-person paradigm is particularly striking, as the extant literature has reported differential infant behaviour when infants are involved in a social interaction and when they act as an observer of other people outside of interactional contexts (Beier & Spelke, 2012; Floor & Akhtar, 2006; Herold & Akhtar, 2008; Moll et al., 2007; Moll & Tomasello, 2007). This suggests that infant cognition may work differently when they are in the 'second-person' position as compared to when they are in the 'third-person' position. Therefore, to better understand the neural mechanisms underpinning infants' active engagement in a social situation, the second-person paradigm is essential in investigations using neural measures.

When transitioning to the second-person paradigm, laying reliable groundwork is critical. To explore the neural mechanisms underlying an infant-adult social interaction, it is optimal to investigate the dynamic relationship that can occur between interacting individuals' brain activities collected simultaneously in a social interaction ('simultaneous dual-brain approach'; Redcay & Schilbach, 2019). A small number of studies have been conducted taking this simultaneous dual-brain approach with infant participants, and have shown that infants' and adults' brain activities are temporally aligned via social cues (Leong et al., 2017; Santamaria et al., 2020). It has also been shown that there is brain-behaviour coordination between an infant and their parent during a social interaction, whereby parental brain activity patterns change according to shifts in infants' looking behaviour (Wass et al., 2018). Whilst these studies are important in showing the feasibility of the approach with infant-adult dyad participants, their analyses focused on just two electrodes. Whilst limiting the analysis to the data from two electrodes was to avoid confounds from motor artefact, this has made it difficult to locate the source of the observed neural activity and infer the underlying cognition. Hence, the function of the observed brain-to-brain or brain-to-behaviour coordination in infant social cognition is yet to be fully understood (Leong et al., 2017; Santamaria et al., 2020; Wass, Noreika, et al., 2018). To better understand this, it is optimal to first make use of a relatively controlled paradigm to identify relevant neural markers that can be expected to emerge in a social interactional context, and then examine the identified markers in a less controlled and more dynamic setting. Naturalistic social interactional environments are rich in information in multiple modalities. When there are rich and complex perceptual inputs and behaviours, it is difficult to understand what drives a certain neural activity identified in the data. Controlled experimental paradigms are effective in establishing reliable neural markers of observable social behaviour, and this is an essential step when transitioning towards less controlled dynamic settings where participants can engage in naturalistic interactions.

The current thesis aimed to build the groundwork for infant social cognition research taking the second-person cognitive neuroscience approach, using a more dynamic social paradigm that involves naturalistic social interactions. The key objectives of this thesis were twofold: first, to establish and validate a protocol for data acquisition and analysis for research using electroencephalogram (EEG) in a more naturalistic and dynamic setting, where infants are actively engaging in a social interaction with another person. Second, to use this technique to explore the interplay between infant cognition and the social world surrounding them. By conducting five studies utilising various paradigms, from a meta-analysis, a conventional screen-based setting, to a dynamic live social interactional situation, we demonstrated that it is feasible to conduct infant social cognition research taking the second-person cognitive neuroscience approach with electrophysiological measures. In the following sections of this chapter, we discuss how such research can advance our understanding about infant social cognition, while identifying and addressing methodological challenges to tackle.

Summary of findings

Evaluating the current status of the second-person dual-brain approach

Given the importance of using more dynamic interactional paradigms to better understand the neural mechanisms that underlie reciprocal social interactions, an increasing number of studies with adults have been conducted taking the second-person simultaneous dual-brain approach using ‘hyperscanning’ techniques, whereby two individuals’ brain activities are monitored simultaneously during a joint activity. Yet, a significant challenge still to tackle in hyperscanning research is to differentiate between the inter-brain connectivity (i.e., common brain activity patterns between individuals) simply resulting from the common perceptual inputs from the shared experimental environment and the inter-brain connectivity that might emerge due to the social interaction in which participants are engaging (Burgess, 2013; Hamilton, 2021). Considering the large heterogeneity in methodological designs and analytical protocols across studies (Czeszumski et al., 2020), a systematic meta-analytic review would help to compare studies and assess any potential methodological artefacts.

In Chapter 2, we conducted a meta-analysis examining the existing research utilising a hyperscanning paradigm reporting inter-brain connectivity occurring in a social situation. The analysis reviewed 12 fMRI studies (reporting 80 results), 24 EEG studies (132 results) and 57 fNIRS studies (230 results) and identified a medium to large level of variations in reported effect sizes. Importantly, our analysis found that the heterogeneity observed in effect sizes was largely associated with differences across studies, rather than different experimental designs or analytic methods. This suggests that the variance of effect sizes reported is likely to result from high researchers' degrees of freedom in choosing research designs and making analytic decisions in each specific study, such as data preprocessing methods and inclusion/exclusion criteria (Botvinik-Nezer et al., 2020; Cohen, 2017; Gemignani & Gervain, 2021; Silberzahn et al., 2018; Stets & Reid, 2011). The results emphasise the need to improve the comparability across different studies to advance our knowledge about the neural mechanisms that underpin reciprocal social interactions. This includes the need to establish a standardised protocol for hyperscanning research, as well as to engage in open and accountable research practices.

Building a stable platform for second-person research utilising a conventional paradigm

When attempting to establish reliable methods that can be applied in the second-person cognitive neuroscience research involving infant participants, a conventional screen-based paradigm helps to identify the specific relationships of interest between neural markers and social behaviour. Beyond the aim of transitioning towards second-person paradigms, conventional 'third-person' paradigms can be informative methods and shed light on the interplay between infant cognition and specified socio-environmental factors.

In Chapter 3, we presented a study that utilised a conventional screen-based paradigm as an example of 'groundwork' studies for the second-person cognitive neuroscience research of infant social cognition. The study explored how infants differently encode information from two adults who give gaze cues to a target object with different levels of accuracy. The literature has documented that infants as young as 8 months of age are already selective learners, and can modify their looking behaviour according to the reliability of others as an informant (Tummeltshammer et al., 2014). Yet, the online process of how infants track such informant reliability, namely how the information about other people's reliability is encoded, had not been investigated. Aiming to address the question, we monitored 9-month-old infants' brain activity whilst they were presented with two informants, one reliable (i.e., always gives a congruent cue) and the other unreliable (i.e., gives an incongruent cue 75% of the time). Utilising the event-related potential (ERP) technique, we found infants' differential neural responses to the two informants, as reflected in the amplitude of the Negative Central component (Nc), as well as the Positive Slow Wave (PSW). These suggest differential levels of infant attentional engagement and differential degrees of information encoding according to the informant reliability (de Haan, 2013; Hoehl et al., 2012; Kopp & Lindenberger, 2011; Parise et al., 2008; Reid et al., 2004; Striano et al., 2006). The evidence from this study extends the existing behavioural reports by shedding light on the neural and cognitive mechanisms of infants' selective learning and behaviour. Importantly, this study shows how infant cognition is modulated by socio-environmental factors, such as another person's gaze cues, pointing to the importance of studying infant social cognition in a social context.

Introducing more dynamic stimuli

Moving towards a more dynamic paradigm from a screen-based procedure, Chapter 4 presented a study where we explored the neural correlates of infants' processing of unexpected events using dynamic video stimuli. Violations of expectations (VoE) have been typically utilised as a paradigm to explore infants' knowledge about the physical and social world (Baillargeon, 2004). However, recently, it has been proposed that VoE might facilitate infants' learning (Köster et al., 2020; Stahl & Feigenson, 2015). Nonetheless, the cognitive processes underlying VoE were yet to be fully understood. Whilst a few studies proposed some neural correlates of VoE (Berger, Tzur, & Posner, 2006; Kaufman, Csibra, & Johnson, 2003, 2005; Köster, Langeloh, & Hoehl, 2019; Köster, Langeloh, Michel, & Hoehl, 2021; Reid et al., 2009), these studies typically focused on one type of VoE event, making it challenging to argue the identified neural correlates reflect the generic processing of VoE rather than the processing of that specific event. To assess whether these neural correlates were markers of general or specific VoE, we presented 8-month-old infants with dynamic video presentations that depicted seven different types of VoE events. Overall, the data did not show clear ERP or event-related oscillation (ERO) effects. Whilst some waveform morphologies might indicate ERP components, the question needs to be readdressed with a modified design to obtain more robust and reliable effects. These results were likely to be related to the use of multiple dynamic videos, which might have caused latency jittering across stimuli (Cohen, 2014). Handling latency jitter is a challenge of using dynamic perceptual inputs, but it may also be an indication of infants' differential processing of different VoE events. Therefore, it is possible that different neural markers index different types of VoE.

Implementing the second-person cognitive neuroscience approach moving beyond a screen-based presentation

To examine the feasibility and potential of conducting infant electrophysiological research using a live interactional setting, applying the second-person cognitive neuroscience approach, in Chapter 5, we explored infant cognition underlying word learning that occurs in social interactions. Learning novel words is an important component of infancy, and research has shown that social interactions play a critical role in infant language learning (Hakuno et al., 2017; Kuhl, 2007a; Kuhl et al., 2003). Yet, typically, research investigating the neural process of word learning utilise screen-based stimuli, removing the social aspects of word learning (e.g., Friedrich & Friederici, 2008, 2011). The study we presented in Chapter 5 investigated the neural online processes of 10-month-old infants underlying their encoding of novel lexical semantic information (i.e., words) presented in a naturalistic live social interaction. We also explored the trajectory from encoding to consolidation of the knowledge, given the existing evidence was mixed regarding how infant lexical knowledge is established as long-term memories, and it was assumed that social interactions are critical to build a robust memory structure (Friedrich & Friederici, 2008, 2011; Friedrich et al., 2015; Henderson et al., 2013).

Analysing neural oscillatory activity, we found that the alpha-band (6-9Hz) as well as the upper bound of alpha-band oscillatory activity (9-10Hz) increased from the first block to the second block of the learning interaction phase. This could be interpreted as attention attenuation and neural habituation (Colombo & Mitchell, 2009; Grill-Spector et al., 2006), suggesting infants have encoded the presented information within a few minutes, despite this task occurring in a socially richer context and thus potentially more cognitively demanding than a conventional experimental setup. Furthermore, we identified individual differences in word learning performance, in terms of the degree of semantic information encoding and attentional engagement to the task. Importantly, these individual differences were found to influence the level of knowledge consolidation assessed after the learning event with and without a delay. This study showed the feasibility of the second-person neuroscience approach with infant participants. Furthermore, this study advanced our knowledge about infant word learning by demonstrating that infants need offline, slow consolidation to acquire semantic lexical knowledge, and that their domain-general social-cognitive skills, including attention and information encoding in a social situation, modulate the trajectory from the encoding of semantic lexical information to its consolidation as knowledge.

Applying the second-person cognitive neuroscience approach in a traditional behavioural paradigm involving a social interaction

In Chapter 6, we presented a study where we trialled the second-person cognitive neuroscience approach in a traditional behavioural paradigm to explore the neural and cognitive processes underlying infant behaviour. We aimed to identify differences in neural activities between accurate and inaccurate search performances in the A-not-B search task reported by Piaget (1954). Piaget's 'A-not-B error' (or 'perseverative error') is one of the most replicated infant behaviours, but there is still little consensus as to when it starts and why it happens. This is likely related to the use of different procedural manipulations that may highlight different cognitive components associated with this paradigm (Dunn & Bremner, 2019). Therefore, using another measure in addition to conventional behavioural measures was considered optimal to explore different aspects of infant cognition underlying the perseverative error without having to modify the traditional task procedure (Dunn & Bremner, 2019). Neural measures can be useful as they can directly examine potential differences in infant cognition between when they err and when they do not. Importantly, the investigation needs to focus on infants' neural activities occurring while infants actively engage in a social task with an experimenter, because it has been shown that infants' search is affected by socio-environmental factors (Dunn & Bremner, 2019; Topál et al., 2008). Prior studies utilising neural measures to investigate infant cognition during this task typically utilised a screen-based task or measured baseline neural activity before the search task is conducted. Therefore, it is important to establish a protocol taking the 'second-person' cognitive neuroscience approach that could be implemented in this traditional Piagetian search task without sacrificing its social nature.

We therefore sought to identify systematic differences in neural activity occurring before accurate and inaccurate searches of 9-month-old infants who engaged in the traditional procedure of Piagetian perseverative search task in a live manner. Due to an issue identified with the setup, the current sample did not make the perseverative error as expected from previous studies. Nevertheless, differences in the neural activities between accurate and inaccurate searches were examined to provide a proof of concept for research using EEG investigating infants' perseverative error. Due to the small sample size and large individual differences observed, no robust neural differences between accurate and inaccurate searches were identified. Yet, to our knowledge, this study provided the first proof of concept that demonstrated the feasibility of utilising electrophysiological measures in this traditional interactive behavioural paradigm of Piagetian perseverative search task, in such a way that it does not interfere with its standard procedure.

Methodological implications

The first key objective of this thesis was to explore a feasible and optimal way of conducting research with infant participants taking the second-person neuroscience approach, where participants' brain activity is monitored when they actively engage in a social interactional situation rather than being passive observers. Our meta-analytic review presented in Chapter 2 has highlighted that we need reliable groundwork to make credible inferences about inter-brain coupling that emerges during a social interaction. Towards that goal, it is important to design research informed by prior evidence demonstrating the relationship between specific brain activity and social interactional elements. Therefore, we moved on to exploring different paradigms that could provide a foundation for future research using hyperscanning techniques to identify the neural mechanisms underlying a reciprocal interaction between an infant and an adult. Our endeavour consisted of two steps. First, we examined how conventional research paradigms involving screen-based presentations could inform future research that utilises a live interactive paradigm (Chapter 3 and 4). Second, we explored the feasibility of studying infants' brain activity while they interact with others in a live manner (Chapter 5 and 6). With these studies, this thesis proposed several experimental paradigms to identify an interpretable and reliable biomarker of social cognition, positioning us closer to the goal of conducting research which uses a naturalistic interactional paradigm taking the second-person cognitive neuroscience approach with hyperscanning techniques.

Whilst research taking the third-person approach involving a screen-based stimuli presentation does not directly investigate the neural mechanisms of reciprocal infant-adult social interactions, it can lay important groundwork for the second-person research that can address the question. Indeed, our study presented in Chapter 3 successfully identified the neural markers of infants' selective learning, and demonstrated how infant cognition interacts with social contextual factors. The neural markers identified in this study could be used in future research investigating how social contexts might affect the working of infant cognition using a live and dynamic setting.

At the same time, work presented in this thesis highlighted challenges of conducting a study using dynamic stimuli as well as dynamic interactional paradigms, instead of static images presented on a screen. In Chapter 4, we were faced with a high level of data attrition due to artefacts from eye movement and gross motor movement. This likely resulted from the use of dynamic video stimuli. A high level of data attrition was also seen in Chapter 6, when a naturalistic and interactional paradigm was used. From such a dynamic paradigm, it is unlikely that we could obtain as many usable datasets as we could from a controlled laboratory setting, as more datasets are likely to be contaminated by bodily movement (Noreika et al., 2020). This reiterates the importance of identifying neural markers of social behaviour that inform an *a priori* hypotheses before conducting research in a dynamic setting. Combined with an advanced method of motor artefact cleaning (e.g., Georgieva et al., 2020), a data-driven hypothesis would help ameliorate the problem of low signal-to-noise ratio due to the data attrition, if not fully resolve the issue. Furthermore, our data in Chapter 4, where we used multiple-event presentations, suggested jittering in the latency of neural responses to violations of expectation events. This is also likely to be due to the use of multiple dynamic events, as some events might have been more salient and easier for infants to process than others. There are advantages of using multiple types of stimuli, such as to maintain infants' attentional engagement in the task to ensure we obtain a sufficient number of trials to conduct a reliable analysis (Stets et al., 2013). Yet, the latency jittering resulting from the processing of various perceptual inputs is a challenge which we need to address when we are aiming to move towards an unconstrained structure of interaction involving rich and multimodal information. This further emphasises the importance of starting with a more controlled setting and gradually loosening the control, so we know what neural markers are associated with infants' specific cognition and behaviour, even when multiple and diverse stimuli are involved.

Another way to address the challenge of utilising dynamic stimuli and settings is the use of an oscillation analysis, also known as a frequency analysis. Oscillatory rhythms in EEG data are considered to reflect the large networks of neurons firing together (Hoehl, Michel, et al., 2014; Jones et al., 2015). Whilst the spatial resolution of EEG is poorer compared to other methods such as fMRI, the magnitude of EEG oscillatory activity as well as its scalp location can represent the selectivity and sensitivity of brain regions and indicate cognitive processes engaged during a certain task (Jones et al., 2015).

Importantly, a power-based analysis such as a frequency analysis is much more robust to latency jittering than a phase-based analysis including ERPs and EROs (Cohen, 2014). Given these advantages, the use of EEG oscillation analysis is becoming more common in research involving infant participants, especially when a live presentation of stimuli is incorporated in the design (e.g., Hoehl, Michel, Reid, Parise, & Striano, 2014; Jones et al., 2015; Orekhova, Stroganova, & Posikera, 1999; Orekhova, Stroganova, Posikera, & Elam, 2006; Reid, Striano, & Iacoboni, 2011). The studies included in this thesis have also demonstrated that an oscillatory analysis can be used in a naturalistic interactive setting where infants engage in an interaction with others rather than act as passive observers of social stimuli presented on a monitor (Chapter 5, 6, also see Discussion in Chapter 4), and that these analyses can provide reliable evidence (Chapter 5). A disadvantage of using a frequency analysis might be a relative lack of prior evidence on which to base interpretations of the data, compared to phase-based analysis, with ERPs being most commonly used in infant research using cognitive neuroscience methods. Yet, there has been an attempt to investigate the association between ERPs and oscillatory activities in terms of their functional characteristics (e.g., Kaduk, 2016; Köster et al., 2021; also Chapter 4 in this thesis). This line of inquiry will be beneficial for the field, as it would enable us to compare the knowledge from ERP studies and the evidence from studies using a dynamic setting in which precise temporal analysis might not be optimal and an oscillatory analysis is more suitable.

Beyond these, further collective effort as a field is necessary to build a stable ground for research taking the second-person cognitive neuroscience approach to investigate the working of infant social cognition in a dynamic social environment. As discussed in Chapter 2, open and transparent research practices and collaborations could inform and facilitate the standardisation of research procedure. It includes, for instance, using clear operational definitions and sharing a reproducible study protocol (Silberzahn et al., 2018).

Theoretical implications

How infant cognition is situated in a social environment

The work included in this thesis demonstrated how infants' attentional and cognitive processes are modulated by other people's behaviour, highlighting the importance of not separating infants from social environmental contexts when studying the functional development of infant cognition. The issue has been discussed in Chapter 3, taking infants' and young children's selective behaviour as an example, where we demonstrated how infants' cognition is modulated by other people's gaze congruency, or their reliability as an informant. Importantly, the results of this study provide a means to consider such an effect when discussing infants' cognition which emerges with the presence of others. For instance, as has been discussed in Chapter 6, infant behaviour in a cognitive task taking place in a social format is shown to be affected by an interaction with an experimenter (Dunn & Bremner, 2019; Topál et al., 2008). The neural markers identified in our study (Chapter 3), including the Nc and PSW ERP components, would make it possible to investigate infants' knowledge about the experimenter and how it might influence infants' behaviour at a cognitive behavioural task.

With the aim of moving towards more dynamic and less controlled experimental settings, future investigations could modify the paradigm used in the study presented in Chapter 3 into a live format. Whilst we were not able to conduct the live version of this study for this thesis due to government restrictions in response to the coronavirus pandemic in 2020 and 2021, the results of the study shown in Chapter 3 would be useful for such future research, as it provided the neural markers of infants' discrimination between different types of social partners in terms of their gaze cueing behaviour. To our knowledge, much of the research investigating the neural correlates of eye gaze processing to date typically has utilised a standard 'third-person' 'isolation' paradigm, where infants are presented with screen-based stimuli depicting other people's face or eyes without experiencing reciprocal interactions with others (e.g., Farroni et al., 2002; Hoehl et al., 2009). Given the evidence suggesting that infant behaviour differs depending on whether they are involved in an interaction or not (Moll et al., 2007; Moll & Tomasello, 2007), it would benefit the field to re-examine what has been suggested from the third-person paradigms using a second-person paradigm which involves a live social interaction, as there may be differences in the working of infant social cognition depending on the perspective they take in a social situation.

Based on the study protocols and paradigms presented in this thesis, it could also be possible to develop a naturalistic live paradigm where infants can engage in a reciprocal gaze exchange with their interaction partner. That line of research could greatly advance our understanding on infant cognitive, socioemotional and language development. Infants' skills to coordinate their attention according to their partners' gaze (e.g., joint attention skills) have been linked to the enhanced functioning of various cognitive skills during infancy, including attention and memory (e.g., Farroni, Massaccesi, Menon, & Johnson, 2007; Hood, Macrae, Cole-Davies, & Dias, 2003; Wass et al., 2018), as well as other developmental outcomes at later stages of postnatal life, such as imitative learning, theory of mind (i.e., understandings about other people's mental state) and language acquisition (Brooks & Meltzoff, 2015; Hobson, 2002; Tomasello, 1999). Yet, we do not yet fully understand the neural mechanisms of such links. Our study presented in Chapter 5 demonstrated an example of connecting neural data from a learning event and a learning outcome, to provide insights into the trajectory of infant learning and cognitive development. Applying this approach, it would be interesting to examine how infants' neural activity during, for instance, live joint attentional episodes would be associated with the trajectory of cognitive, socioemotional and language development. Furthermore, understanding neural correlates of infants' processing of gaze in a live interaction is an essential step towards dynamic, dual-brain second-person research which aims to shed light on the neural mechanisms of gaze exchanges between an infant and an adult (Leong et al., 2017; Wass et al., 2018). By building a stable platform that leads to a dual-brain second-person cognitive neuroscience research, it would be possible to address the challenges of the existing second-person research and explore the function of inter-brain connectivity that emerges between an infant and an adult during exchanges of social cues (Leong et al., 2017; Santamaria et al., 2020; Wass, Noreika, et al., 2018).

Towards the integration of different social behaviour during infancy

Investigating the neural correlates of infants' social referencing induced by violations of expectation

The study presented in Chapter 4 aimed to identify the neural correlates of infants' violations of expectation (VoE). This study could serve as a foundation for future investigations looking into infants' social behaviour, beyond the processing of unexpected events. Infants' ability to detect and process VoE has been shown to be a precursor of social active learning (Dunn & Bremner, 2017). It has also been proposed that developing the abilities to process unexpectancy might mark key milestones in the social and cognitive development during infancy (Köster et al., 2020; Stahl & Feigenson, 2019). These suggest an importance and potential of specifying neural markers of infants' VoE to better understand infant social cognitive development.

The link between VoE and infant learning can be seen in the evidence showing that VoE leads to social referencing (Dunn & Bremner, 2017). Social referencing refers to infant looking behaviour whereby they look to adults, primarily their parent, when they face a novel, uncertain or unexpected situation (Field, Diego, & Hernandez-Reif, 2009; Forman, Minick, & Stone, 1996; Rochat, 2014; Tronick, Als, & Adamson, 1979; Walden, Kim, McCoy, & Karrass, 2007; Walden & Ogan, 1988). As infants seek comfort and reassurance under conditions of uncertainty, social referencing has often been associated with emotional bonding and attachment between infants and parents (Ainsworth, 1992; Dickstein, Thompson, Estes, Malkin, & Lamb, 1984). Yet, evidence suggests that it is better interpreted as information seeking behaviour rather than comfort seeking (Bazhydai et al., 2020; Dunn & Bremner, 2017; Stenberg, 2009). By 9 months of age, infants look to an unfamiliar experimenter as much as they look to the primary caregiver when they are faced with an ambiguous situation, indicating that social referencing can happen to people with whom infants do not have secure attachment (Kutsuki et al., 2007). At around 12 months of age, infants choose whom to look to based on the context at hand, rather than simply turning to a primary caregiver (Bazhydai et al., 2020; Stenberg, 2009). These reports suggest that social referencing occurs when infants seek information from others, expecting their looking behaviour to elicit social informative cues from a potential informant. Given the function of social referencing as infants' information seeking in a social situation, it is possible to interpret social referencing more generally as a way for infants to initiate a social interaction with others via gazing behaviour. Therefore, understanding the neural substrates of social referencing would indicate what neural and cognitive processes underlie infants' active learning in a social situation.

When we identify the neural activity that emerges prior to social referencing triggered by infants' VoE, that neural activity is likely to include both the processing of VoE and the neural 'trigger' of social referencing. Hence, it is critical to identify the neural correlates of VoE before investigating the neural process of infants' social referencing. Whilst our study presented in Chapter 4 could not specify the neural substrates of VoE due to the latency jittering across different event stimuli, we believe that the study has provided a proof of concept, and validated a paradigm and protocol for future research, which aims to investigate the neural correlates of infants' social referencing induced by VoE.

Infants' social cognition and detection of unexpectancy

Beyond VoE and social referencing, it has recently been proposed that a number of developmental phenomena could be described under the framework of predictive processing model, according to which VoE plays a critical role to trigger and enhance infant learning (Köster et al., 2020; Stahl & Feigenson, 2019). As introduced in Chapter 4, the predictive processing approach views human cognition as a computational system, whereby predictions are made based on perceptual inputs and produce optimal behaviour (Clark, 2013; Schubotz, 2015). Whilst our study presented in Chapter 4 focused on infants' VoE regarding non-social object movements, the existing evidence shows that infants can make predictions and detect its violation in various other domains including social action sequences (e.g., Reid et al., 2009) and lexical information (e.g., Parise & Csibra, 2012). This suggests that, if the neural correlates of general VoE are identified, it could shed light on the cognitive processes underlying infants' understanding of the social world.

An intriguing example of infants' VoE in the social domain is how young children learn to understand other people's actions and the underlying intentions, including joking and pretence. Studies have investigated when young children start to understand people's joking or pretence behaviour, and indicated that when action sequences are familiar enough for young children to predict the action consequence, they are able to discriminate different intentions behind the same or similar actions (Behne, Carpenter, Call, & Tomasello, 2005; Hoicka & Gattis, 2008; Onishi, Baillargeon, & Leslie, 2007). For instance, Hoicka and Gattis (2008) examined toddlers' understanding of ambiguous actions which can be interpreted either as a joke or a mistake (e.g., bringing a spoon up to one's face and hitting their cheek). They showed that even 24-month-old children have difficulty understanding other people's joking actions (Hoicka & Gattis, 2008). In their study, ambiguous actions were marked by either positive (jokes) or negative (mistakes) affect by the demonstrator of the action. Ambiguous actions were defined as 'incomplete actions that could be interpreted either as jokes or mistakes (e.g., putting a hat on over one's eyes). The result showed that 19- to 24-month-olds could not discriminate ambiguous humour from an ambiguous mistake, as the frequency of them imitating the actions did not differ across conditions. However, when the action itself was not ambiguous (i.e., the action was dissimilar to actions normally applied to the used object: e.g., putting a boot on one's head), they were able to discriminate jokes from mistakes. These results suggest that the detection of unconventional, or unexpected, action sequences may play a critical role in infants' understanding another person's intention to joke.

In a similar vein, Onishi, Baillargeon and Leslie (2007) investigated how 15-month-old toddlers understand other people's pretence play, utilising a VoE paradigm. The authors compared whether toddlers would show differential looking responses to the experimenter's pretending to drink when familiar and unfamiliar objects were used (cups, shoes and tubes). The experimenter *pretended to* pour drink into one of the objects presented, and they subsequently pretended to drink from the object they pretended to pour the drink into (expected), or the object they did not pour (unexpected). The analysis on the looking time suggested that toddlers could differentiate between expected and unexpected sequences when cups were used, but not when shoes or tubes were used. Importantly, when a short familiarisation phase was introduced where infants saw an actor drinking from a shoe or a tube, they were able to differentiate an expected sequence from an unexpected one. The examination of the effect of the familiarisation phase suggests that the novelty or unfamiliarity of the action hinders toddlers' successful processing of pretence actions. These findings further support the idea that infants' knowledge about what to be expected plays a critical role in their understanding of other people's intentions.

In sum, understanding the neural markers and cognitive mechanisms of infants' VoE could shed light on how infants come to understand complex social elements such as other people's intentions and become capable of engaging in various social situations including joking and play. Pretend play, or children's understanding of pretence, has been considered to be a critical developmental milestone, contributing to the development of cognitive, socioemotional and language development (Fein, 1981; Lillard et al., 2012). Therefore, investigating how children learn to understand other people's different action intentions via VoE may also shed light on other aspects of child development in cognitive, socioemotional and language domains.

To that end, the importance of conducting research taking the second-person cognitive neuroscience approach persists. It is interesting to notice that the evidence mentioned above about toddlers' understanding of other people' actions was typically provided from studies utilising the third-person paradigm, where toddlers act as passive observers of others. Whilst these 'third-person' studies suggest that toddlers aged 24 months struggle to infer other people’s intention to joke, when infants can take the second-person perspective, evidence indicates infants as young as 9-months can understand complex action intentions. For instance, Behne and colleagues (2005) demonstrated that 9- and 12-month-olds showed different behavioural responses to other people’s mistakes and teasing. In this study, infants were faced with an experimenter who did not give a toy to infants in a teasing fashion (i.e., pull back the toy, smiling, when infants try to reach it) or because they were unable to do so (i.e., try to give the toy to the infant but accidentally drop it) (Behne et al., 2005). When the experimenter was teasing, infants showed behaviour that indicated their frustration, such as more reaching, banging or longer time of looking away, whereas when the experimenter accidentally dropped a toy and was unable to pass it to infants, they responded more patiently, with less reaching, banging or looking away (Behne et al., 2005). Whilst the infants’ behavioural differences do not necessarily mean that they have an understanding of people’s psychological states behind different behaviours, the results do suggest that 9-month-old infants can differentiate the situations where people are and are not willing to give them a toy. Yet, whilst the evidence might look contradictory, these reported results referred to above are not entirely comparable in terms of their design. For example, alongside the difference in the second- and third-person positions of infants’, the differences in the experimenter’s actions across conditions in Hoicka and Gattis (2008) are much more subtle than Behne et al. (2005). In the study by Hoicka and Gattis (2008), the experimenter demonstrated the same action but showed different emotional expressions (laughter for the joking condition, “woops” for the mistake condition), whereas the experimenter’s behaviour in Behne et al. (2005) differ to a larger degree across conditions. Furthermore, Hoicka and Gattis (2008) measured infants’ imitative responses, which is much more demanding than looking and banging that were assessed by Behne et al. (2005). Therefore, it might be the sensitivity of measures used as well as the saliency of behavioural cues that differentiated the conclusions of the two studies (Bremner and Dunn, 2020). That said, it is certainly possible that the critical factor might be whether the infant was in the second- or third-person position in relation to the experimenter. The teasing condition in Behne et al. (2005) was much more interactive, involving contingent and reciprocal exchanges between the experimenter and toddler, as the experimenter had to pull the toy when the toddler reached for the toy. Therefore, experimental conditions where infants can actively participate in an interaction might have provided infants with information about other people’s actions and underlying intentions which would not have been available if they were outside of the interaction as observers (Butterfill, 2013).

Future studies are expected to explore these questions, and we believe that our studies presented in the thesis provided a template for research that could lead to such investigation.

Overall, identifying neural markers of infants' VoE to better understand the underlying cognitive processes can inform future research investigating the neural and cognitive mechanisms of social referencing, and other studies exploring the development of infants' understanding of other people's action intentions. We believe that the study presented in Chapter 4 of this thesis has positioned us a step closer towards that goal.

Potential of the second-person cognitive neuroscience approach in infant social cognition research

The work presented in this thesis not only highlighted the importance of research taking the second-person approach but also showed the feasibility of conducting such research investigating infant cognition in a live interactional setting, where infants are able to interact with another person in a naturalistic and contingent manner. Despite the challenges of utilising dynamic settings as highlighted in Chapter 4, in Chapters 5 and 6, this thesis provided a proof of concept for research investigating infant social cognition in a live interactional setting, utilising a single-brain second-person cognitive neuroscience approach (Chapters 5 and 6). Importantly, we demonstrated that it is possible to do so without altering standardised procedures of conventional behavioural tasks (Chapter 6).

Research taking this approach can shed light on the neural mechanisms of infant learning and cognitive development that occurs in the social environment. The neural correlates of infant cognition and learning have often been explored in a context 'isolated' from social situations (Chapter 1). Yet, as our work presented in Chapters 3 and 5 demonstrated, attentional and cognitive processes are modulated by social environmental factors. Furthermore, individual differences in attending to complex social situations affect infants' learning (Chapter 5), which coincide with the results shown in Chapter 3 regarding infants' cognition underlying infants' selective social behaviour. Therefore, we must continue to conduct research taking this second-person approach, which investigates infants' cognition *within* a social context. As this thesis demonstrated, revisiting neural evidence gained through the 'isolation' third-person paradigm from the second-person perspective can advance our understanding on the working of infant cognition in a social situation. Further, such research taking the second-person approach could inform the designs of future research utilising hyperscanning techniques to explore the neural mechanisms that underlie reciprocal infant-adult interactions, and how infant leaning and development occur in the dynamic social environment.

Reflection and directions for future research

Our work presented in this thesis has focused on developing protocols for research taking the second-person cognitive neuroscience methods to explore infant social cognition in a social setting, as well as investigating how socio-environmental contexts affect infant cognition. Our discussion was centred on how we could advance our understanding by incorporating neurophysiological measures into a dynamic paradigm. This endeavour started with the idea that infant cognition needs to be studied in as naturalistic a context as possible while maintaining the rigour of experimental procedure to conduct a robust and interpretable analysis. We believe this goal has been achieved as we have shown what is currently feasible, and proposed how the foundational work we have presented could lead to future investigations.

Nevertheless, our work is not free from limitations. First, we could not include any studies that utilised hyperscanning techniques to address the question of what neural mechanisms underlie infant-adult social interactions. This was largely due to practical reasons, including the time constraints of my doctoral research funding, as well as the global pandemic that occurred in 2020 and 2021, which made in-person data acquisitions impossible. If not for the global pandemic, for instance, the study presented in Chapter 4 investigating the neural correlates of VoE would have led to another study aiming to achieve the same goal with a modified paradigm. We could also have examined the neural substrates of social referencing in infants based on the results, leading to the investigation of brain-to-brain coordination between an infant and an adult to whom the infant looks to, utilising hyperscanning techniques. The study in Chapter 6 could also be re-conducted to examine systematic differences in neural activities between infants who do and do not make the Piagetian perseverative errors, by addressing the setup issue identified. Subsequently, we could examine the relationship between an infant's and an experimenter's brain activities as an index of the social aspect of this search task. Whilst these adapted and extended versions of these studies could have taken place, we believe this thesis has laid informative, reliable, and promising foundational work for future research when the situation allows.

Second, the work presented in this thesis did not directly compare infant cognition between when they are in the second-person position and when they are in the third-person position. This is due to the design of the studies presented in this thesis, of which the central aim was not the comparison between these perspectives. To our knowledge, there has been no neurophysiological study that explored the relevant question, whilst some behavioural studies provided evidence that the working of infant cognition differs depending on infant's perspectives (Moll et al., 2007; Moll & Tomasello, 2007). Hence, it is hoped that these questions be addressed by future research.

Relatedly, whilst the narrower definition of social cognition refers to our understanding of other people's minds, the studies included in this thesis did not directly investigate the very question of how infants understand other people's mental states. Instead, in this thesis, we focused on how social interactional settings might modulate infant cognition and learning (Chapters 5 and 6). The methods we proposed and verified in this thesis would be applicable to research that aims to investigate the question, and we believe the approach whose usability this thesis demonstrated would be of great use. Taking infants' understanding of other people's action intentions for example, the research in this field typically utilises stimuli isolated from a social component (e.g., arm movement only with the actor's face not shown), due to the typical conceptualisation of action being perceptual stimuli (Nyström, Ljunghammar, Rosander, & Von Hofsten, 2011; van Elk et al., 2008). Yet, the working of the mirror neuron systems, which has been associated with action perception (Nyström et al., 2011), has been shown to be modulated by social contextual factors (Meyer et al., 2021). Another body of research investigated how action is interpreted semantically (Cummings, Čeponiene, Dick, Saygin, & Townsend, 2008; Pace, Carver, & Friend, 2013; Reid et al., 2009; Reid & Striano, 2008). It has been shown that infants make use of situational knowledge to understand the semantics of other people's action sequences (Ní Choisdealbha & Reid, 2014). This, again, points to the importance of studying infants' action understanding without separating them from a social context. Future research could address these questions taking the second-person approach and a naturalistic dynamic paradigm, such as ones presented in this thesis (Chapters 5 and 6).

On this topic of social cognition, a body of studies which suggest the early emergence of theory of mind, or understanding about other people's mental states (e.g., Kampis, Parise, Csibra, & Kovács, 2015; Kovács et al., 2010) has mainly utilised conventional third-person paradigms, instead of second-person interactive paradigms. In these studies, infants are presented with stimuli in which another person observes an event from a different perspective from that of the infant's (Kampis et al., 2015; Kovács et al., 2010). Both neural and behavioural studies have suggested infants can represent the other person's perspective (Kampis et al., 2015; Kovács et al., 2010). This indicates that, unlike a claim made by behavioural studies by Moll and colleagues (2007) that infants can only understand other people's mental experiences from the second-person position, infants younger than 12 months of age might have a rudimentary concept of other people's mental states that are different from their own (Kampis et al., 2015; Kovács et al., 2010). The mixed evidence is intriguing and might be due to the difference in used measures, rather than the difference in infant perspectives. Neural measures have been particularly effective for young infants with limited behavioural repertoire, as neural measures could detect differences in perceptual and cognitive processes even in the absence of behavioural indication (Parise et al., 2008). It would be interesting to examine whether infants might show a sign of other people's perspectives when they are in an interaction with others whose mental states they have to infer, and how social communicative cues, such as direct eye contact and infant-directed speech might modulate their social understanding (Moll et al., 2007).

Third, we were not able to consider a wider social context, such as culture and ethnicity. The participants in our studies mainly came from families who were white and in middle-class population, and no more precise demographic information was collected or analysed due to no specific rationale to do so based on prior research. These limit the generalisability of our results. Given that social interactional conventions differ across cultures (van Ijzendoorn & Sagi-Schwartz, 2008), it is possible that infant social cognition might be modulated by wider cultural contexts as well as the given social situation at hand. It would be interesting to examine cross-cultural differences in infant social cognition that emerges during a social interaction.

Last, but not least, understanding the developmental changes in infant social cognition is important, yet this has not been addressed in the current thesis. The developmental trajectory of infant social behaviour has been well documented; for instance, infants transition from a dyadic to a triadic relationship at around the age of 9 months (Rochat, 2001). Protodeclarative pointing, as an example of infant active engagement to a social partner, emerges between 11 and 12 months of age, whereas protoimperative pointing starts between 12 and 13 months (Carpenter et al., 1998). These demonstrate how infant social behaviour develops and becomes more sophisticated with age. Longitudinal study designs would enable the tracking of developmental changes in social cognition underlying social behaviour as well as its neural correlates. Importantly, the second-person cognitive neuroscience approach would allow us to investigate such changes in a paradigm with high ecological validity, provided that it is feasible to link a certain behaviour with a specific neural activity even when a naturalistic and dynamic interactional paradigm is used, as shown in the Chapters 5 and 6 of this thesis. We hope that future research would incorporate longitudinal designs to examine the developmental trajectory of social cognition from the second-person perspective during infancy.

Along with the suggestions made in the previous sections, returning to our starting point, further effort is necessary to build a stable foundation for research of infant social cognition taking the second-person cognitive neuroscience approach with a paradigm where infants can actively interact with other people. These efforts could lead to a standardisation of paradigms and measures of such research, which is critical to conduct reliable analyses which can yield more credible knowledge.

Conclusion

The research conducted as part of this thesis has shown that it is feasible to conduct studies which investigate infant social cognition taking the second-person cognitive neuroscience approach, where infant brain activity is monitored while the infant actively takes part in a social interaction with others. The work presented in this thesis showed how studies using screen-based stimuli could identify reliable neural markers, which is essential as a ‘base model’ when moving towards the second-person research involving a more naturalistic and dynamic setting. Furthermore, we demonstrated that this second-person cognitive neuroscience approach can be applied to traditional and standardised behavioural study procedures, which allows us to investigate the neural and cognitive mechanisms of infant social behaviour that occurs in a live social situation. Results indicated that the working of infants’ cognition is indeed embedded in social contexts, reiterating the importance of conducting infant social cognition research using a paradigm where infants are not isolated from a social interactional situation. These studies jointly demonstrate the potentials of conducting infant social cognition research taking the second-person cognitive neuroscience approach to better understand how infants’ learning and cognition are linked to their social position, and their ability to process their surrounding social environment. It is hoped that our work presented in this thesis has inspired future research that explores how infant social cognition works and develops, interacting with the dynamic social environment.

Consolidated bibliography

Ackerman, J. M., & Kenrick, D. T. (2009). Cooperative courtship: Helping friends raise and raze relationship barriers. *Personality and Social Psychology Bulletin*, *35*(10), 1285–1300. https://doi.org/10.1177/0146167209335640

Ackles, P. K., & Cook, K. G. (1998). Stimulus probability and event-related potentials of the brain in 6-month-old human infants: A parametric study. *International Journal of Psychophysiology*, *29*(2), 115-143. https://doi.org/10.1016/S0167-8760(98)00013-0

Ackles, P. K., & Cook, K. G. (2007). Attention or memory? Effects of familiarity and novelty on the Nc component of event-related brain potentials in six-month-old infants. *International Journal of Neuroscience*, *117*(6), 837-867. https://doi.org/10.1080/00207450600909970

Ainsworth, M. D. S. (1992). A consideration of social referencing in the context of attachment theory and research. In S. Feinman (Ed.), *Social referencing and the social construction of reality in infancy* (pp. 349–367). Springer. https://doi.org/10.1007/978-1-4899-2462-9_14

Altvater-Mackensen, N., & Grossmann, T. (2015). Learning to match auditory and visual speech cues: Social influences on acquisition of phonological categories. *Child Development*, *86*(2), 362–378. https://doi.org/10.1111/cdev.12320

Amoruso, L., Gelormini, C., Aboitiz, F., Alvarez González, M., Manes, F., Cardona, J., & Ibanez, A. (2013). N400 ERPs for actions: Building meaning in context. *Frontiers in Human Neuroscience, 7*, 57. https://doi.org/10.3389/fnhum.2013.00057

Asch, S. E. (1952). *Social psychology.* Prentice-Hall Inc.

Astolfi, L., Cincotti, F., Mattia, D., De Vico Fallani, F., Salinari, S., Vecchiato, G., … Babiloni, F. (2010). Imaging the social brain: multi-subjects EEG recordings during the “Chicken’s game”. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society,* 1734–1737. https://doi.org/10.1109/IEMBS.2010.5626708

Astolfi, L, Toppi, J., Ciaramidaro, A., Vogel, P., Freitag, C. M., & Siniatchkin, M. (2020). Raising the bar: Can dual scanning improve our understanding of joint action? *NeuroImage*, *216*, 116813. https://doi.org/10.1016/j.neuroimage.2020.116813

Ayrolles, A., Brun, F., Chen, P., Djalovski, A., Beauxis, Y., Delorme, R., … Dumas, G. (2021). HyPyP: a hyperscanning Python pipeline for inter-brain connectivity analysis. *Social Cognitive and Affective Neuroscience*, *16*(1–2), 72–83. https://doi.org/10.1093/scan/nsaa141

Azhari, A., Truzzi, A., Neoh, M. J. Y., Balagtas, J. P. M., Tan, H. A. H., Goh, P. L. P., … Esposito, G. (2020). A decade of infant neuroimaging research: What have we learned and where are we going? *Infant Behavior and Development*, *58*, 101389. https://doi.org/10.1016/j.infbeh.2019.101389

Babiloni, F., & Astolfi, L. (2014). Social neuroscience and hyperscanning techniques: Past, present and future. *Neuroscience and Biobehavioral Reviews*, *44*, 76–93. https://doi.org/10.1016/j.neubiorev.2012.07.006

Babiloni, F., Astolfi, L., Cincotti, F., Mattia, D., Tocci, A., Tarantino, A., … De Vico Fallani, F. (2007). Cortical activity and connectivity of human brain during the prisoner's dilemma: an EEG hyperscanning study. *IEEE Engineering in Medicine and Biology Society. Annual Conference*, 4953–4956. https://doi.org/10.1109/IEMBS.2007.4353452

Babiloni, F., Tocci, A., De Vico Fallani, F., Cincotti, F., Mattia, D., Bianchi, L., … Astolfi, L. (2007). High resolution EEG hyperscanning during a card game. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society,* 4957–4960. https://doi.org/10.1109/IEMBS.2007.4353453

Baccalá, L. A., & Sameshima, K. (2001). Partial directed coherence: a new concept in neural structure determination. *Biological Cybernetics, 84*(6), 463–474. https://doi.org/10.1007/PL00007990

Baillargeon, R. (1987). Object permanence in 3 1/2- and 4 1/2-month-old infants. *Developmental Psychology*, *23*(5), 655–664. https://doi.org/10.1037/0012-1649.23.5.655

Baillargeon, R. (1987). Young infants' reasoning about the physical and spatial properties of a hidden object. *Cognitive Development*, *2*(3), 179-200. https://doi.org/10.1016/S0885-2014(87)90043-8

Baillargeon, R. (2004). Infants' physical world. *Current Directions in Psychological Science*, *13*(3), 89–94. https://doi.org/10.1111/j.0963-7214.2004.00281.x

Baillargeon, R, & DeVos, J. (1991). Object permanence in young infants: Further evidence. *Child Development*, *62*(6), 1227–1246. https://doi.org/10.1111/j.1467-8624.1991.tb01602.x

Baillargeon, R, DeVos, J., & Graber, M. (1989). Location memory in 8-month-old infants in a non-search AB task: Further evidence. *Cognitive Development*, *4*(4), 345–367. https://doi.org/10.1016/S0885-2014(89)90040-3

Baillargeon, R, & Graber, M. (1987). Where's the rabbit? 5.5-month-old infants' representation of the height of a hidden object. *Cognitive Development*, *2*(4), 375–392. https://doi.org/10.1016/S0885-2014(87)80014-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

Baillargeon, R, & Wang, S. (2002). Event categorization in infancy. *Trends in Cognitive Sciences*, *6*(2), 85–93. https://doi.org/10.1016/S1364-6613(00)01836-2

Baird, A. A., Kagan, J., Gaudette, T., Walz, K. A., Hershlag, N., & Boas, D. A. (2002). Frontal lobe activation during object permanence: Data from near-infrared spectroscopy. *NeuroImage*, *16*(4), 1120–1126. https://doi.org/10.1006/nimg.2002.1170

Baker, J. M., Liu, N., Cui, X., Vrticka, P., Saggar, M., Hosseini, S. M. H. H., & Reiss, A. L. (2016). Corrigendum: Sex differences in neural and behavioral signatures of cooperation revealed by fNIRS hyperscanning. *Scientific Reports*, *6*, 30512. https://doi.org/10.1038/srep30512

Bakker, M., Sommerville, J. A., & Gredebäck, G. (2016). Enhanced neural processing of goal-directed actions after active training in 4-month-old infants. *Journal of Cognitive Neuroscience*, 28(3), 472-482. https://doi.org/10.1162/jocn_a_00909

Balliet, D., Li, N. P., Macfarlan, S. J., & Van Vugt, M. (2011). Sex differences in cooperation: A meta-analytic review of social dilemmas. *Psychological Bulletin*, *137*(6), 881–909. https://doi.org/10.1037/a0025354

Baron-Cohen, S. (1989). Perceptual role taking and protodeclarative pointing in autism. *British Journal of Developmental Psychology*, *7*(2), 113–127. https://doi.org/10.1111/J.2044-835X.1989.TB00793.X

Baron-Cohen, S. (1995). *Mindblindness—an essay on autism and theory of mind*. MIT Press. https://doi.org/10.7551/mitpress/4635.001.0001

Barr, R., Muentener, P., Garcia, A., Fujimoto, M., & Chávez, V. (2007). The effect of repetition on imitation from television during infancy. *Developmental Psychobiology*, *49*(2), 196–207. https://doi.org/10.1002/dev.20208

Barraza, P., Dumas, G., Liu, H., Blanco-Gomez, G., van den Heuvel, M. I., Baart, M., & Pérez, A. (2019). Implementing EEG hyperscanning setups. *MethodsX*, *6*, 428–436. https://doi.org/10.1016/j.mex.2019.02.021

Başar, E., & Güntekin, B. (2012). A short review of alpha activity in cognitive processes and in cognitive impairment. *International Journal of Psychophysiology*, *86*(1), 25-38. https://doi.org/10.1016/j.ijpsycho.2012.07.001

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Batki, A., Baron-Cohen, S., Wheelwright, S., Connellan, J., & Ahluwalia, J. (2000). Is there an innate gaze module? Evidence from human neonates. *Infant Behavior and Development*, *23*(2), 223-229. https://doi.org/10.1016/S0163-6383(01)00037-6

Bazhydai, M., Westermann, G., & Parise, E. (2020). “I don’t know but I know who to ask”: 12-month-olds actively seek information from knowledgeable adults. *Developmental Science*, *23*(5), e12938. https://doi.org/10.1111/desc.12938

Becchio, C., Bertone, C., & Castiello, U. (2008). How the gaze of others influences object processing. *Trends in cognitive sciences*, *12*(7), 254-258. https://doi.org/10.1016/j.tics.2008.04.005

Becchio, C., Sartori, L., & Castiello, U. (2010). Toward you: The social side of actions. *Current Directions in Psychological Science*, *19*(3), 183–188. https://doi.org/10.1177/0963721410370131

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

Begus, K., Gliga, T., & Southgate, V. (2016). Infants’ preferences for native speakers are associated with an expectation of information. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(44), 12397–12402. https://doi.org/10.1073/pnas.1603261113

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). https://doi.org/10.1098/rsbl.2015.0041

Behne, T., Carpenter, M., Call, J., & Tomasello, M. (2005). Unwilling versus unable: Infants' understanding of intentional action. *Developmental Psychology*, *41*(2), 328–337. https://doi.org/10.1037/0012-1649.41.2.328

Beier, J. S., & Spelke, E. S. (2012). Infants' developing understanding of social gaze. *Child Development*, *83*(2). https://doi.org/10.1111/J.1467-8624.2011.01702.X

Bell, M. A. (2002). Power changes in infant EEG frequency bands during a spatial working memory task. *Psychophysiology*, *39*(4), 450-458. https://doi.org/10.1017.S0048577201393174

Bell, M. A., & Fox, N. A. (1992). The relations between frontal brain electrical activity and cognitive development during infancy. *Child Development*, *63*(5), 1142–1163. https://doi.org/10.1111/j.1467-8624.1992.tb01685.x

Bell, M. A., & Fox, N. A. (1997). Individual differences in object permanence performance at 8 months: Locomotor experience and brain electrical activity. *Developmental Psychobiology*, *31*(4), 287–297. https://doi.org/10.1002/(SICI)1098-2302(199712)31:4<287::AID-DEV6>3.0.CO;2-N

Benavides-Varela, S., Gómez, D. M., Macagno, F., Bion, R. A. H., Peretz, I., & Mehler, J. (2011). Memory in the neonate brain. *PLoS ONE*, *6*(11), e27497. https://doi.org/10.1371/journal.pone.0027497

Benninger, C., Matthis, P., & Scheffner, D. (1984). EEG development of healthy boys and girls. Results of a longitudinal study. *Electroencephalography and Clinical Neurophysiology*, *57*(1), 1-12. https://doi.org/10.1016/0013-4694(84)90002-6

Berger, A., Tzur, G., & Posner, M. I. (2006). Infant brains detect arithmetic errors. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(33), 12649–12653. https://doi.org/10.1073/pnas.0605350103

Bickhard, M. H. (2009). Interactivism: A manifesto. *New Ideas in Psychology*, *27*(1), 85–95. https://doi.org/10.1016/j.newideapsych.2008.05.001

Bickhard, M. H. (2016). Inter- and en-activism: Some thoughts and comparisons. *New Ideas in Psychology*, *41*, 23–32. https://doi.org/10.1016/j.newideapsych.2015.12.002

Bloom, P. (2004). Myths of word learning. In D. G. Hall & S. R. Waxman (Eds.), *Weaving a lexicon* (pp. 205–224). MIT Press. https://doi.org/10.7551/mitpress/7185.001.0001

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis* (2nd ed.). Wiley.

Bosseler, A. N., Taulu, S., Pihko, E., Mäkelä, J. P., Imada, T., Ahonen, A., & Kuhl, P. K. (2013). Theta brain rhythms index perceptual narrowing in infant speech perception. *Frontiers in Psychology*, *4*, 690. https://doi.org/10.3389/fpsyg.2013.00690

Botvinik-Nezer, R., Holzmeister, F., Camerer, C. F., Dreber, A., Huber, J., Johannesson, M., … Schonberg, T. (2020). Variability in the analysis of a single neuroimaging dataset by many teams. *Nature*, *582*(7810), 84–88. https://doi.org/10.1038/s41586-020-2314-9

Bouchon, C., Nazzi, T., & Gervain, J. (2015). Hemispheric asymmetries in repetition enhancement and suppression effects in the newborn brain. *PLoS ONE*, *10*(10), e0140160. https://doi.org/10.1371/journal.pone.0140160

Boyer, T. W., Samantha Pan, J., & Bertenthal, B. I. (2011). Infants' understanding of actions performed by mechanical devices. *Cognition*, *121*(1), 1–11. https://doi.org/10.1016/j.cognition.2011.05.012

Bremner, A., & Bryant, P. (2001). The effect of spatial cues on infants' responses in the AB task, with and without a hidden object. *Developmental Science*, *4*(4), 408–415. https://doi.org/10.1111/1467-7687.00182

Bremner, J. G., & Bryant, P. E. (1977). Place versus response as the basis of spatial errors made by young infants. *Journal of Experimental Child Psychology*, *23*(1), 162–171. https://doi.org/10.1016/0022-0965(77)90082-0

Bremner, J. G., & Dunn, K. (2020). What's in a look? How can we best measure infants' response to incongruity?. *Human Development*, *64*(2), 64-67. https://doi.org/10.1159/000510579

Bremner, J. G., & Knowles, L. S. (1984). Piagetian Stage IV search errors with an object that is directly accessible both visually and manually. *Perception*, *13*(3), 307–314. https://doi.org/10.1068/p130307

Bremner, J. G. (1985). Object tracking and search in infancy: A review of data and a theoretical evaluation. *Developmental Review*, *5*(4), 371–396. https://doi.org/10.1016/0273-2297(85)90018-8

Bremner, J. G. (1998). Commentary on Munakata. *Developmental Science*, *1*(2), 192–193. https://doi.org/10.1111/1467-7687.00024

Brettle, A. J., & Long, A. F. (2001). Comparison of bibliographic databases for information on the rehabilitation of people with severe mental illness. *Bulletin of the Medical Library Association*, *89*(4), 353–362.

Brooks, R., & Meltzoff, A. N. (2004). Gaze following: A mechanism for building social connections between infants and adults. In M. Mikulincer & P. R. Shaver (Eds.), *Mechanisms of social connection: From brain to group.* (pp. 167–183). American Psychological Association. https://doi.org/10.1037/14250-010

Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science*, *8*(6), 535–543. https://doi.org/10.1111/j.1467-7687.2005.00445.x

Brooks, R., & Meltzoff, A. N. (2015). Connecting the dots from infancy to childhood: A longitudinal study connecting gaze following, language, and explicit theory of mind. *Journal of Experimental Child Psychology*, *130*. https://doi.org/10.1016/j.jecp.2014.09.010

Brown, H., Weighall, A., Henderson, L. M., & Gaskell, M. G. (2012). Enhanced recognition and recall of new words in 7- and 12-year-olds following a period of offline consolidation. *Journal of Experimental Child Psychology*, *112*(1), 56–72. https://doi.org/10.1016/j.jecp.2011.11.010

Brown, V. A. (2021). An introduction to linear mixed-effects modeling in R. *Advances in Methods and Practices in Psychological Science*, *4*(1). https://doi.org/10.1177/2515245920960351

Bruner, J. S. (1974). From communication to language — a psychological perspective. *Cognition*, *3*(3), 255-287. https://doi.org/10.1016/0010-0277(74)90012-2

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

Bürkner, P.C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Bürkner, P.C. (2018). Advanced Bayesian multilevel modeling with the R Package brms. *The R Journal*, *10*(1), 395–411. https://doi.org/10.32614/RJ-2018-017

Buttelmann, D., Zmyj, N., Daum, M., & Carpenter, M. (2013). Selective imitation of in-group over out-group members in 14-month-old infants. *Child Development*, *84*(2), 422–428. https://doi.org/10.1111/j.1467-8624.2012.01860.x

Butterfill, S. A. (2013). Interacting mindreaders. *Philosophical Studies*, *165*(3), 841–863. https://doi.org/10.1007/s11098-012-9980-x

Butterworth, G. (1975). Object identity in infancy: The interaction of spatial location codes in determining search errors. *Child Development*, *46*(4), 866–870. https://doi.org/10.2307/1128391

Butterworth, G. (1977). Object disappearance and error in Piaget's Stage IV task. *Journal of Experimental Child Psychology*, *23*(3), 391–401. https://doi.org/10.1016/0022-0965(77)90034-0

Butterworth, G., & Jarrett, N. (1991). What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology*, *9*(1), 55–72. https://doi.org/10.1111/j.2044-835x.1991.tb00862.x

Cannon, E. N., Woodward, A. L., Gredebäck, G., von Hofsten, C., & Turek, C. (2012). Action production influences 12-month-old infants' attention to others' actions. *Developmental Science*, *15*(1), 35–42. https://doi.org/10.1111/j.1467-7687.2011.01095.x

Carey, S. (2010). Beyond fast mapping. *Language Learning and Development*, *6*(3), 184–205. https://doi.org/10.1080/15475441.2010.484379

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, *63*(4), 1–174. https://doi.org/10.2307/1166214

Cashon, C. H., & Cohen, L. B. (2000). Eight-month-old Infants' perception of possible and impossible events. *Infancy*, *1*(4), 429–446. https://doi.org/10.1207/S15327078IN0104_4

Cheng, X., Li, X., & Hu, Y. (2015). Synchronous brain activity during cooperative exchange depends on gender of partner: A fNIRS-based hyperscanning study. *Human Brain Mapping*, *36*(6), 2039–2048. https://doi.org/10.1002/hbm.22754

Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Näätänen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nature Neuroscience*, *1*, 351–353. https://doi.org/10.1038/1561

Cheung, M. W. L. (2019). A guide to conducting a meta-analysis with non-independent effect sizes. *Neuropsychology Review, 29*(4), 387-396. https://doi.org/10.1007/s11065-019-09415-6

Chita-Tegmark, M. (2016). Social attention in ASD: A review and meta-analysis of eye-tracking studies. *Research in Developmental Disabilities*, *48*, 79–93. https://doi.org/10.1016/j.ridd.2015.10.011

Choudhury, N., & Gorman, K. S. (2000). The relationship between sustained attention and cognitive performance in 17-24-month old toddlers. *Infant and Child Development*, *9*(3), 127–146. https://doi.org/10.1002/1522-7219(200009)9:3<127::AID-ICD225>3.0.CO;2-5

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181-204. https://doi.org/10.1017/S0140525X12000477

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. W., Diedrich, F. J., Smith, L. B., & Thelen, E. (2006). Young infants reach correctly in A-not-B tasks: On the development of stability and perseveration. *Infant Behavior and Development*, *29*(3), 435–444. https://doi.org/10.1016/j.infbeh.2006.03.001

Clearfield, M. W., Dineva, E., Smith, L. B., Diedrich, F. J., & Thelen, E. (2009). Cue salience and infant perseverative reaching: Tests of the dynamic field theory. *Developmental Science*, *12*(1), 26–40. https://doi.org/10.1111/j.1467-7687.2008.00769.x

Clément, F. (2010). To trust or not to trust? Children's social epistemology. *Review of Philosophy and Psychology, 1*(4), 531-549. https://doi.org/10.1007/s13164-010-0022-3

Cleveland, A., & Striano, T. (2008). Televised social interaction and object learning in 14- and 18-month-old infants. *Infant Behavior and Development*, *31*(2), 326–331. https://doi.org/10.1016/j.infbeh.2007.12.019

Cohen, L. B. (1977). Habituation of infant visual attention. In T. J. Tighe & R. N. Leaton (Eds.), *Habituation: Perspectives from child development, animal behavior, and neurophysiology* (pp. 207–238). Lawrence Erlbaum.

Cohen, M. X. (2017). Rigor and replication in time-frequency analyses of cognitive electrophysiology data. *International Journal of Psychophysiology*, *111*, 80–87. https://doi.org/10.1016/j.ijpsycho.2016.02.001

Cohen, M. X. (2011). It's about time. *Frontiers in Human Neuroscience*, 1–16. https://doi.org/10.3389/fnhum.2011.00002

Cohen, M. X. (2014). *Analyzing neural time series data: Theory and practice*. MIT Press. https://doi.org/10.1017/CBO9781107415324.004

Cohn, J. F., & Tronick, E. Z. (1988). Mother-infant face-to-face interaction: Influence is bidirectional and unrelated to periodic cycles in either partner's behavior. *Developmental Psychology*, *24*(3), 386–392. https://doi.org/10.1037/0012-1649.24.3.386

Colombo, J., & Mitchell, D. W. (2009). Infant visual habituation. *Neurobiology of Learning and Memory*, *92*(2), 225–234. https://doi.org/10.1016/j.nlm.2008.06.002

Conboy, B. T., Brooks, R., Meltzoff, A. N., & Kuhl, P. K. (2015). Social interaction in infants' learning of second-language phonetics: An exploration of brain-behavior relations. *Developmental Neuropsychology*, *40*(4), 216–229. https://doi.org/10.1080/87565641.2015.1014487

Costall, A. (2006). "Introspectionism" and the mythical origins of scientific psychology. *Consciousness and Cognition*, *15*(4), 634–654. https://doi.org/10.1016/j.concog.2006.09.008

Costall, A., Leudar, I., & Reddy, V. (2006). Failing to see the irony in "mind-reading," *16*(2), 163–167. https://doi.org/10.1177/0959354306062533

Courchesne, E., Ganz, L., & Norcia, A. M. (1981). Event-related brain potentials to human faces in infants. *Child Development*, 804-811. https://doi.org/10.1111/j.1467-8624.1981.tb03117.x

Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & Language*, *25*(2), 141–168. https://doi.org/10.1111/j.1468-0017.2009.01384.x

Csibra, G., & Gergely, G. (2006a). Social learning and social cognition: The case for pedagogy. In Y. Munakata & M. H. Johnson (Eds.), *Processes of change in brain and cognitive development. Attention and performance, XXI* (pp. 249–274). Oxford University Press.

Csibra, G., & Gergely, G. (2006b). Sylvia's recipe: The role of imitation and pedagogy in the transmission of cultural knowledge. In N. J. Enfield & S. C. Levenson (Eds.), *Roots of human sociality: Culture, cognition, and human interaction* (pp. 229–255). Routledge. https://doi.org/10.4324/9781003135517

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

Csibra, G., & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1567), 1149-1157. https://doi.org/10.1098/rstb.2010.0319

Cuevas, K., Cannon, E. N., Yoo, K., & Fox, N. A. (2014). The infant EEG mu rhythm: Methodological considerations and best practices. *Developmental Review*, *34*(1), 26-43. https://doi.org/10.1016/j.dr.2013.12.001

Cuevas, K., Swingler, M. M., Bell, M. A., Marcovitch, S., & Calkins, S. D. (2012). Measures of frontal functioning and the emergence of inhibitory control processes at 10 months of age. *Developmental Cognitive Neuroscience*, *2*(2), 235–243. https://doi.org/10.1016/j.dcn.2012.01.002

Cui, X., Bryant, D. M., & Reiss, A. L. (2012). NIRS-based hyperscanning reveals increased interpersonal coherence in superior frontal cortex during cooperation. *NeuroImage*, *59*(3), 2430–2437. https://doi.org/10.1016/j.neuroimage.2011.09.003

Cummings, A., Čeponiene, R., Dick, F., Saygin, A. P., & Townsend, J. (2008). A developmental ERP study of verbal and non-verbal semantic processing. *Brain Research*, *1208*, 137–149. https://doi.org/10.1016/j.brainres.2008.02.015

Cummings, E. M., & Bjork, E. L. (1981). The search behavior of 12 to 14 month-old infants on a five-choice invisible displacement hiding task. *Infant Behavior and Development*, *4*(1), 47–60. https://doi.org/10.1016/S0163-6383(81)80007-0

Cummings, E. M., & Bjork, E. L. (1983). Perseveration and search on a five-choice visible displacement hiding task. *Journal of Genetic Psychology*, *142*(2), 283–291. https://doi.org/10.1080/00221325.1983.10533519

Czeszumski, A., Eustergerling, S., Lang, A., Menrath, D., Gerstenberger, M., Schuberth, S., … König, P. (2020). Hyperscanning: A valid method to study neural inter-brain underpinnings of social interaction. *Frontiers in Human Neuroscience*, 14, 39. https://doi.org/10.3389/fnhum.2020.00039

D'Entremont, B., Hains, S. M. J., & Muir, D. W. (1997). A demonstration of gaze following in 3- to 6-month-olds. *Infant Behavior and Development*, *20*(4), 569–572. https://doi.org/10.1016/S0163-6383(97)90048-5

Daum, M. M., Prinz, W., & Aschersleben, G. (2011). Perception and production of object-related grasping in 6-month-olds. *Journal of Experimental Child Psychology*, *108*(4), 810–818. https://doi.org/10.1016/j.jecp.2010.10.003

Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning: Neural and behavioural evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1536), 3773-3800. https://doi.org/10.1098/rstb.2009.0111

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

de Haan, M. (2013). *Infant EEG and event-related potentials*. Psychology.

de Haan, M., Belsky, J., Reid, V., Volein, A., & Johnson, M. H. (2004). Maternal personality and infants' neural and visual responsivity to facial expressions of emotion. *Journal of Child Psychology and Psychiatry*, *45*(7), 1209-1218. https://doi.org/10.1111/j.1469-7610.2004.00320.x

de Haan, M., & Gunnar, M. R. (2009). *Handbook of developmental social neuroscience*. Guilford Press.

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

de Haan, M., & Nelson, C. A. (1997). Recognition of the mother's face by six-month-old infants: A neurobehavioral study. *Child Development*, *68*(2), 187-210. https://doi.org/10.1111/j.1467-8624.1997.tb01935.x

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

De Jaegher, H., & Di Paolo, E. A. (2007). Participatory sense-making: An enactive approach to social cognition. *Phenomenology and the Cognitive Sciences*, *6*(4), 485–507. https://doi.org/10.1007/S11097-007-9076-9

De Jaegher, H., & Di Paolo, E. A. (2013). Enactivism is not interactionism. *Frontiers in Human Neuroscience*, *6*, 345. https://doi.org/10.3389/FNHUM.2012.00345

Deák, G. O., Flom, R. A., & Pick, A. D. (2000). Effects of gesture and target on 12- and 18-month-olds' joint visual attention to objects in front of or behind them. *Developmental Psychology*, *36*(4), 511–523. https://doi.org/10.1037//0012-1649.36.4.511

DeBoer, T., Scott, L., & Nelson, C. A. (2004). Event-related potentials in developmental populations. In T. C. Handy (Ed.), *Event-related potentials: A methods handbook* (pp. 263–297). MIT Press.

Diamond, A. (1988). Abilities and neural mechanisms underlying AB performance. *Child Development*, *59*(2), 523–527. https://doi.org/10.1111/j.1467-8624.1988.tb01486.x

Diamond, A. (1985). Development of the ability to use recall to guide action, as indicated by infants' performance on AB. *Child Development*, *56*(4), 868–883. https://doi.org/10.1111/j.1467-8624.1985.tb00160.x

Dickstein, S., Thompson, R. A., Estes, D., Malkin, C., & Lamb, M. E. (1984). Social referencing and the security of attachment. *Infant Behavior and Development*, *7*(4), 507–516. https://doi.org/10.1016/S0163-6383(84)80009-0

Diedrich, F. J., Highlands, T. M., Spahr, K. A., Thelen, E., & Smith, L. B. (2001). The role of target distinctiveness in infant perseverative reaching. *Journal of Experimental Child Psychology*, *78*(3), 263–290. https://doi.org/10.1006/jecp.2000.2569

Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., … Poeppel, D. (2017). Brain-to-brain synchrony tracks real-world dynamic group interactions in the classroom. *Current Biology*, *27*(9), 1375–1380. https://doi.org/10.1016/j.cub.2017.04.002

Dineva, E., & Schöner, G. (2018). How infants' reaches reveal principles of sensorimotor decision making. *Connection Science*, *30*(1), 53–80. https://doi.org/10.1080/09540091.2017.1405382

Dockrell, J. E., Braisby, N., & Best, R. M. (2007). Children's acquisition of science terms: Simple exposure is insufficient. *Learning and Instruction*, *17*(6), 577–594. https://doi.org/10.1016/j.learninstruc.2007.09.005

Dommer, L., Jäger, N., Scholkmann, F., Wolf, M., & Holper, L. (2012). Between-brain coherence during joint n-back task performance: A two-person functional near-infrared spectroscopy study. *Behavioural Brain Research*, *234*(2), 212–222. https://doi.org/10.1016/j.bbr.2012.06.024

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition*, *6*(5), 509-540. https://doi.org/10.1080/135062899394920

Dumas, G., Nadel, J., Soussignan, R., Martinerie, J., & Garnero, L. (2010). Inter-brain synchronization during social interaction. *PloS One*, *5*(8), e12166. https://doi.org/10.1371/journal.pone.0012166

Duncan J, Owen AM. Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences.* 2000; 23:475–483. https://doi.org/10.1016/s0166-2236(00)01633-7

Dunn, K., & Bremner, J. G. (2017). Investigating looking and social looking measures as an index of infant violation of expectation. *Developmental Science*, *20*(6). https://doi.org/10.1111/desc.12452

Dunn, K., & Bremner, J. G. (2019). Investigating the social environment of the A-not-B search task. *Developmental Science*, *23*(3), e12921. https://doi.org/10.1111/desc.12921

Eigsti, I. M., De Marchena, A. B., Schuh, J. M., & Kelley, E. (2011). Language acquisition in autism spectrum disorders: A developmental review. *Research in Autism Spectrum Disorders*, *5*(2), 681–691. https://doi.org/10.1016/j.rasd.2010.09.001

Elsner, B., & Pauen, S. (2007). Social learning of artefact function in 12- and 15-month-olds. *European Journal of Developmental Psychology*, *4*(1), 80–99. https://doi.org/10.1080/17405620601051220

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

Farroni, T, Mansfield, E. M., Lai, C., & Johnson, M. H. (2003). Infants perceiving and acting on the eyes: Tests of an evolutionary hypothesis. *Journal of Experimental Child Psychology*, *85*(3), 199–212. https://doi.org/10.1016/S0022-0965(03)00022-5

Farroni, T, Massaccesi, S., Menon, E., & Johnson, M. H. (2007). Direct gaze modulates face recognition in young infants. *Cognition*, 102(3), 396-404. https://doi.org/10.1016/j.cognition.2006.01.007

Fein, G. G. (1981). Pretend play in childhood: An integrative review. *Child Development*, *52*(4), 1095. https://doi.org/10.2307/1129497

Feinman, S., Roberts, D., Hsieh, K.F., Sawyer, D., & Swanson, D. (1992). A critical review of social referencing in infancy. *Social Referencing and the Social Construction of Reality in Infancy*, 15–54. https://doi.org/10.1007/978-1-4899-2462-9_2

Feldman, R. (2007). Parent-infant synchrony and the construction of shared timing; physiological precursors, developmental outcomes, and risk conditions. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *48*(3–4), 329–354. https://doi.org/10.1111/j.1469-7610.2006.01701.x

FeldmanHall, O., & Shenhav, A. (2019). Resolving uncertainty in a social world. *Nature Human Behaviour*, 3(5), 426-435. https://doi.org/10.1038/s41562-019-0590-x

Feng, C., Thompson, W. K., & Paulus, M. P. (2022). Effect sizes of associations between neuroimaging measures and affective symptoms: A meta-analysis. *Depression and anxiety*, *39*(1), 19-25. https://doi.org/10.1002/da.23215

Field, T. (1978). The three Rs of infant-adult interactions: Rhythms, repertoires, and responsitivity. *Journal of Pediatric Psychology*, *3*(3), 131–136. https://doi.org/10.1093/jpepsy/3.3.131

Field, T., Diego, M., & Hernandez-Reif, M. (2009). Depressed mothers' infants are less responsive to faces and voices. *Infant Behavior and Development*, *32*(3), 239–244. https://doi.org/10.1016/J.INFBEH.2009.03.005

Fishburn, F. A., Murty, V. P., Hlutkowsky, C. O., MacGillivray, C. E., Bemis, L. M., Murphy, M. E., … Perlman, S. B. (2018). Putting our heads together: Interpersonal neural synchronization as a biological mechanism for shared intentionality. *Social Cognitive and Affective Neuroscience*, *13*(8), 841–849. https://doi.org/10.1093/scan/nsy060

Fiske, S. T. (1993). Social cognition and social perception. *Annual Review of Psychology*, *44*, 155–194. https://doi.org/10.1146/annurev.ps.44.020193.001103

Floor, P., & Akhtar, N. (2006). Can 18-month-old infants learn words by listening in on conversations? *Infancy*, *9*(3), 327–339. https://doi.org/10.1207/S15327078IN0903_4

Fogel, A., Diamond, G., Langhorst, B., & Demos, V. (1982). Affective and cognitive aspects of the 2-month-old's participation in face-to-face interaction with the mother. In Edward. Tronick (Ed.), *Social interchange in infancy: affect, cognition, and communication* (pp. 35–57). University Park Press.

Forman, E. A., Minick, N., & Stone, C. A. (1996). *Contexts for learning: sociocultural dynamics in children's development*. Oxford University Press.

Friedrich, M., & Friederici, A. D. (2004). N400-like semantic incongruity effect in 19-month-olds: Processing known words in picture contexts. *Journal of Cognitive Neuroscience*, *16*(8), 1465–1477. https://doi.org/10.1162/0898929042304705

Friedrich, M., & Friederici, A. D. (2005a). Lexical priming and semantic integration reflected in the event-related potential of 14-month-olds. *NeuroReport*, *16*(6), 653–656. https://doi.org/10.1097/00001756-200504250-00028

Friedrich, M., & Friederici, A. D. (2005b). Phonotactic knowledge and lexical-semantic processing in one-year-olds: Brain responses to words and nonsense words in picture contexts. *Journal of Cognitive Neuroscience*, *17*(11), 1785–1802. https://doi.org/10.1162/089892905774589172

Friedrich, M., & Friederici, A. D. (2006). Early N400 development and later language acquisition. *Psychophysiology*, *43*(1), 1–12. https://doi.org/10.1111/j.1469-8986.2006.00381.x

Friedrich, M., & Friederici, A. D. (2008). Neurophysiological correlates of online word learning in 14-month-old infants. *NeuroReport*, *19*(18), 1757–1761. https://doi.org/10.1097/WNR.0b013e328318f014

Friedrich, M., & Friederici, A. D. (2010). Maturing brain mechanisms and developing behavioral language skills. *Brain and Language*, *114*(2), 66–71. https://doi.org/10.1016/j.bandl.2009.07.004

Friedrich, M., & Friederici, A. D. (2011). Word learning in 6-month-olds: Fast encoding-weak retention. *Journal of Cognitive Neuroscience*, *23*(11), 3228-3240. https://doi.org/10.1162/jocn_a_00002

Friedrich, M., & Friederici, A. D. (2017). The origins of word learning: Brain responses of 3-month-olds indicate their rapid association of objects and words. *Developmental Science*, *20*(2), e12357. https://doi.org/10.1111/desc.12357

Friedrich, M., Wilhelm, I., Born, J., & Friederici, A. D. (2015). Generalization of word meanings during infant sleep. *Nature Communications*, *6*(1), 1–9. https://doi.org/10.1038/ncomms7004

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1456), 815–836. https://doi.org/10.1098/rstb.2005.1622

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787

Frith, U., & Frith, C. (2001). The biological basis of social interaction. *Current Direction in Psychological Science*, 151–155. https://doi.org/10.1111/1467-8721.00137

Frith, U., & Frith, C. (2010). The social brain: allowing humans to boldly go where no other species has been. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1537), 165. https://doi.org/10.1098/RSTB.2009.0160

Gallotti, M., & Frith, C. D. (2013). Social cognition in the we-mode. *Trends in Cognitive Sciences*, *17*(4), 160–165. https://doi.org/10.1016/j.tics.2013.02.002

Gemignani, J., & Gervain, J. (2021). Comparing different pre-processing routines for infant fNIRS data. *Developmental Cognitive Neuroscience*, *48*, 100943. https://doi.org/10.1016/j.dcn.2021.100943

Georgieva, S., Lester, S., Noreika, V., Yilmaz, M. N., Wass, S., & Leong, V. (2020). Toward the understanding of topographical and spectral signatures of infant movement artifacts in naturalistic EEG. *Frontiers in Neuroscience*, *14*, 352. https://doi.org/10.3389/fnins.2020.00352

Gergely, G., Bekkering, H., & Király, I. (2002). Rational imitation in preverbal infants. *Nature*, *415*(6873), 755. https://doi.org/10.1038/415755a

Gibson, J. J. (1986). *The ecological approach to visual perception*. Psychology Press.

Goeman, J. J., & Solari, A. (2011). Multiple testing for exploratory research. *Statistical Science*, *26*(4), 584-597. https://doi.org/10.1214/11-STS356

Gopnik, A., & Bonawitz, E. (2015). Bayesian models of child development. *Wiley Interdisciplinary Reviews: Cognitive Science*, *6*(2), 75–86. https://doi.org/10.1002/wcs.1330

Grady, C. L., Maisog, J. M., Horwitz, B., Ungerleider, L. G., Mentis, M. J., Salerno, J. A., ... & Haxby, J. V. (1994). Age-related changes in cortical blood flow activation during visual processing of faces and location. *Journal of Neuroscience*, *14*(3), 1450-1462. https://doi.org/10.1523/jneurosci.14-03-01450.1994

Gratch, G., Appel, K. J., Evans, W. F., LeCompte, G. K., & Wright, N. A. (1974). Piaget's Stage IV object concept error: Evidence of forgetting or object conception? *Child Development*, *45*(1), 71. https://doi.org/10.2307/1127751

Gratch, G., & Landers, W. F. (1971). Stage IV of Piaget's theory of infant's object concepts: A longitudinal study. *Child Development*, *42*(2), 359. https://doi.org/10.2307/1127472

Gredebäck, G., Kaduk, K., Bakker, M., Gottwald, J., Ekberg, T., Elsner, C., … Kenward, B. (2015). The neuropsychology of infants' pro-social preferences. *Developmental Cognitive Neuroscience*, *12*, 106-113. https://doi.org/10.1016/j.dcn.2015.01.006

Gredebäck, G., & Melinder, A. (2011). Teleological reasoning in 4-month-old infants: Pupil dilations and contextual constraints. *PLoS ONE*, *6*(10), 2–6. https://doi.org/10.1371/journal.pone.0026487

Gredebäck, G., Melinder, A., & Daum, M. (2010). The development and neural basis of pointing comprehension. *Social Neuroscience*, *5*(5-6), 441-450. https://doi.org/10.1080/17470910903523327

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, *10*(1), 14-23. https://doi.org/10.1016/j.tics.2005.11.006

Grinsted, A., Moore, J. C., & Jevrejeva, S. (2004). Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Processes in Geophysics*, *11*(5/6), 561–566. https://doi.org/10.5194/NPG-11-561-2004

Grossmann, T. (2015). The development of social brain functions in infancy. *Psychological Bulletin*, *141*(6). https://doi.org/10.1037/bul0000002

Grossmann, T., & Johnson, M. H. (2010). Selective prefrontal cortex responses to joint attention in early infancy. *Biology Letters*, *6*(4), 540–543. https://doi.org/10.1098/rsbl.2009.1069

Grossmann, T., Lloyd-Fox, S., & Johnson, M. H. (2013). Brain responses reveal young infants' sensitivity to when a social partner follows their gaze. *Developmental Cognitive Neuroscience*, *6*, 155–161. https://doi.org/10.1016/j.dcn.2013.09.004

Gruber, M. J., Watrous, A. J., Ekstrom, A. D., Ranganath, C., & Otten, L. J. (2013). Expected reward modulates encoding-related theta activity before an event. *NeuroImage*, *64*, 68–74. https://doi.org/10.1016/j.neuroimage.2012.07.064

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 of the United States of America*, *106*(13), 5365–5370. https://doi.org/10.1073/pnas.0900289106

Gvirts, H. Z., & Perlmutter, R. (2020). What guides us to neurally and behaviorally align with anyone specific? A neurobiological model based on fNIRS hyperscanning studies. *The Neuroscientist*, *26*(2), 108–116. https://doi.org/10.1177/1073858419861912

Hains, S. M. J., & Muir, D. W. (1996). Infant sensitivity to adult eye direction. *Child Development*, *67*(5), 1940–1951. https://doi.org/10.1111/j.1467-8624.1996.tb01836.x

Haith, M. M. (1993). Future-oriented processes in infancy: The case of visual expectations. In C. Granrud (Ed.), *Visual perception and cognition in infancy.* (pp. 235–264). Lawrence Erlbaum Associates Inc.

Hakuno, Y., Omori, T., Yamamoto, J. I., & Minagawa, Y. (2017). Social interaction facilitates word learning in preverbal infants: Word–object mapping and word segmentation. *Infant Behavior and Development*, *48*, 65–77. https://doi.org/10.1016/j.infbeh.2017.05.012

Hamilton, A. F. de C. (2021). Hyperscanning: Beyond the hype. *Neuron*, *109*(3), 404–407. https://doi.org/10.1016/j.neuron.2020.11.008

Hari, R., & Kujala, M. V. (2009). Brain basis of human social interaction: from concepts to brain imaging. *Physiological Reviews*, *89*(2), 453–479. https://doi.org/10.1152/physrev.00041.2007.-Modern

Harrer, M., Cuijpers, P., Furukawa, T., & Ebert, D. D. (2019). dmetar: Companion R package for the guide "Doing meta-analysis in R". R package version 0.0.9000.

Harris, P. L. (1973). Perseverative errors in search by young infants. *Child Development*, *44*(1), 28. https://doi.org/10.2307/1127675

Harris, P. L. (2012). *Trusting what you're told: How children learn from others. Belknap Press.*

Harris, P. L., Koenig, M. A., Corriveau, K. H., & Jaswal, V. K. (2018). Cognitive foundations of learning from testimony. *Annual Review of Psychology*, *69*, 251–273. https://doi.org/10.1146/annurev-psych-122216-011710

Hasson, U., Ghazanfar, A. A., Galantucci, B., Garrod, S., & Keysers, C. (2012). Brain-to-brain coupling: A mechanism for creating and sharing a social world. *Trends in Cognitive Sciences, 16(2), 114-121*. https://doi.org/10.1016/j.tics.2011.12.007

Hasson, U., Landesman, O., Knappmeyer, B., Vallines, I., Rubin, N., & Heeger, D. J. (2008). Neurocinematics: The neuroscience of film. *Projections*, *2*(1), 1–26. https://doi.org/10.3167/proj.2008.020102

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science*, *303*(5664), 1634–1640. https://doi.org/10.1126/science.1089506

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Academic Press.

Henderson, L., Weighall, A., & Gaskell, G. (2013). Learning new vocabulary during childhood: Effects of semantic training on lexical consolidation and integration. *Journal of Experimental Child Psychology*, *116*(3), 572–592. https://doi.org/10.1016/j.jecp.2013.07.004

Herold, K. H., & Akhtar, N. (2008). Imitative learning from a third-party interaction: Relations with self-recognition and perspective-taking. *Journal of Experimental Child Psychology*, *101*(2), 114–123. https://doi.org/10.1016/J.JECP.2008.05.004

Hespos, S. J., & Baillargeon, R. (2001a). Infants' knowledge about occlusion and containment events: A surprising discrepancy. *Psychological Science*, 12(2), 141-147. https://doi.org/10.1111/1467-9280.00324

Hespos, S. J., & Baillargeon, R. (2001b). Reasoning about containment events in very young infants. *Cognition*, 78(3), 207-245. https://doi.org/10.1016/S0010-0277(00)00118-9

Hespos, S. J., & Baillargeon, R. (2008). Young infants' actions reveal their developing knowledge of support variables: Converging evidence for violation-of-expectation findings. *Cognition*, *107*(1), 304-316. https://doi.org/10.1016/j.cognition.2007.07.009

Heyes, C. (2017a). Rattling the cage and opening the door. *Developmental Science*, *20*(2), e12416. https://doi.org/10.1111/DESC.12416

Heyes, C. (2017b). When does social learning become cultural learning? *Developmental Science*, *20*(2), e12350. https://doi.org/10.1111/DESC.12350

Higgins, J. P. T., & Green, S. (2011). Cochrane handbook for systematic reviews of interventions (Version 5.1.0).

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539–1558. https://doi.org/10.1002/sim.1186

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal*, *327*(7414), 557-560. https://doi.org/10.1136/bmj.327.7414.557

Hinde, R. A., & Stevenson-Hinde, J. (1987). Interpersonal relationships and child development. *Developmental Review*, *7*(1), 1–21. https://doi.org/10.1016/0273-2297(87)90002-5

Hinde, S., & Spackman, E. (2014). Bidirectional citation searching to completion: An exploration of literature searching methods. *PharmacoEconomics*, *33*(1), 5–11. https://doi.org/10.1007/s40273-014-0205-3

Hobson, P. (2002). *The cradle of thought: Exploring the origins of thinking*. MacMillan.

Hoehl, S., & Markova, G. (2018). Moving developmental social neuroscience toward a second-person approach. PLoS Biology, *16*(12), e3000055. https://doi.org/10.1371/journal.pbio.3000055

Hoehl, S., Michel, C., Reid, V. M., Parise, E., & Striano, T. (2014). Eye contact during live social interaction modulates infants' oscillatory brain activity. *Social Neuroscience*, *9*(3), 300–308. https://doi.org/10.1080/17470919.2014.884982

Hoehl, S., Reid, V. M., Parise, E., Handl, A., Palumbo, L., & Striano, T. (2009). Looking at eye gaze processing and its neural correlates in infancy - Implications for social development and autism spectrum disorder. *Child Development*, *80*(4), 968–985. https://doi.org/10.1111/j.1467-8624.2009.01311.x

Hoehl, S., Reid, V. M., Mooney, J., & Striano, T. (2008). What are you looking at? Infants' neural processing of an adult's object-directed eye gaze. *Developmental Science*, *11*(1), 10–16. https://doi.org/10.1111/j.1467-7687.2007.00643.x

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

Hoehl, S., Wahl, S., Michel, C., & Striano, T. (2012). Effects of eye gaze cues provided by the caregiver compared to a stranger on infants' object processing. *Developmental Cognitive Neuroscience*, *2*(1), 81–89. https://doi.org/10.1016/j.dcn.2011.07.015

Hoehl, S., Wahl, S., & Pauen, S. (2014). Disentangling the effects of an adult model's eye gaze and head orientation on young infants' processing of a previously attended object. *Infancy, 19(1), 53-64*. https://doi.org/10.1111/infa.12035

Hofstadter, M., & Reznick, J. S. (1996). Response modality affects human infant delayed-response performance. *Child Development*, *67*(2), 646. https://doi.org/10.2307/1131838

Hoicka, E., & Gattis, M. (2008). Do the wrong thing: How toddlers tell a joke from a mistake. *Cognitive Development*, *23*(1), 180–190. https://doi.org/10.1016/J.COGDEV.2007.06.001

Holper, L., Scholkmann, F., & Wolf, M. (2012). Between-brain connectivity during imitation measured by fNIRS. *NeuroImage*, *63*(1), 212–222. https://doi.org/10.1016/j.neuroimage.2012.06.028

Hood, B. M., Macrae, C. N., Cole-Davies, V., & Dias, M. (2003). Eye remember you: The effects of gaze direction on face recognition in children and adults. *Developmental Science, 6(1), 67-71*. https://doi.org/10.1111/1467-7687.00256

Hoormann, J., Falkenstein, M., Schwarzenau, P., & Hohnsbein, J. (1998). Methods for the quantification and statistical testing of ERP differences across conditions. *Behavior Research Methods, Instruments, & Computers, 30*(1), 103–109. https://doi.org/10.3758/BF03209420

Howick, J., Koletsi, D., Ioannidis, J. P., Madigan, C., Pandis, N., Loef, M., ... & Schmidt, S. (2022). Most healthcare interventions tested in Cochrane Reviews are not effective according to high quality evidence: a systematic review and meta-analysis. *Journal of clinical epidemiology*. https://doi.org/10.1016/j.jclinepi.2022.04.017

Hull, C. L. (1943). *Principles of behavior: an introduction to behavior theory*. Appleton-Century.

Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. *Advances in infancy research*, *5*, 69–95.

Hunter, S. J., & Karrer, R. (1993). ERPs indicate event duration effects on infants' visual attention and recognition memory. *Electroencephalography and Clinical Neurophysiology*, *87*(2), S38. https://doi.org/10.1016/0013-4694(93)91000-q

Ilyka, D., Johnson, M. H., & Lloyd-Fox, S. (2021). Infant social interactions and brain development: A systematic review. *Neuroscience & Biobehavioral Reviews*, *130*, 448–469. https://doi.org/10.1016/J.NEUBIOREV.2021.09.001

Jablonka, E. (2015). Information: Its interpretation, its inheritance, and its sharing. *Philosophy of Science*, *69*(4), 578–605. https://doi.org/10.1086/344621

James, E., Gaskell, M. G., Weighall, A., & Henderson, L. (2017). Consolidation of vocabulary during sleep: The rich get richer? *Neuroscience and Biobehavioral Reviews*, *77*, 1–13. https://doi.org/10.1016/j.neubiorev.2017.01.054

Jessor, R. (1958). The problem of reductionism in psychology. *Psychological Review*, *65*(3), 170–178. https://doi.org/10.1037/h0045385

Jiang, J., Dai, B., Peng, D., Zhu, C., Liu, L., & Lu, C. (2012). Neural synchronization during face-to-face communication. *Journal of Neuroscience*, *32*(45), 16064-16069. https://doi.org/10.1523/JNEUROSCI.2926-12.2012

Jones, E. J. H. 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

Kaduk, K. (2016). *The semantic link: action & language: an investigation of relations between different cognitive domains in early development*. [Doctoral thesis, Lancaster University]. Lancaster University Research Directory. https://doi.org/10.17635/lancaster/thesis/174

Kaduk, K., Bakker, M., Juvrud, J., Gredebäck, G., Westermann, G., Lunn, J., & Reid, V. M. (2016). Semantic processing of actions at 9 months is linked to language proficiency at 9 and 18 months. *Journal of Experimental Child Psychology, 151, 96-108*. https://doi.org/10.1016/j.jecp.2016.02.003

Kaduk, K., Elsner, B., & Reid, V. M. (2013). Discrimination of animate and inanimate motion in 9-month-old infants: An ERP study. *Developmental Cognitive Neuroscience*, *6*, 14–22. https://doi.org/10.1016/j.dcn.2013.05.003

Kampe, K. K. W., Frith, C. D., & Frith, U. (2003). “Hey John”: Signals conveying communicative intention toward the self activate brain regions associated with “mentalizing,” regardless of modality. *The Journal of Neuroscience, 23(12), 5258-5263*. https://doi.org/10.1523/jneurosci.23-12-05258.2003

Kampis, D., Parise, E., Csibra, G., & Kovács, Á. M. (2015). Neural signatures for sustaining object representations attributed to others in preverbal human infants. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1819). https://doi.org/10.1098/rspb.2015.1683

Karrer, R., & Ackles, P. K. (1987). Visual event-related potentials of infants during a modified oddball procedure. *Electroencephalography and clinical neurophysiology. Supplement*, *40*, 603–608. Retrieved from https://pubmed.ncbi.nlm.nih.gov/3480181/

Karrer, J. H., Karrer, R., Bloom, D., Chaney, L., & Davis, R. (1998). Event-related brain potentials during an extended visual recognition memory task depict delayed development of cerebral inhibitory processes among 6-month-old infants with Down syndrome. *International Journal of Psychophysiology*, 29(2), 167-200. https://doi.org/10.1016/S0167-8760(98)00015-4

Karrer, R., & Monti, L. A. (1995). Event-related potentials of 4-7-week-old infants in a visual recognition memory task. *Electroencephalography and Clinical Neurophysiology*, *94*(6), 414-424. https://doi.org/10.1016/0013-4694(94)00313-A

Kaufman, J., Csibra, G., & Johnson, M. H. (2003). Representing occluded objects in the human infant brain. *Proceedings of the Royal Society B: Biological Sciences*, *270*(Suppl_2), S140–S143. https://doi.org/10.1098/rsbl.2003.0067

Kaufman, J., Csibra, G., & Johnson, M. H. (2005). Oscillatory activity in the infant brain reflects object maintenance. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(42), 15271–15274. https://doi.org/10.1073/pnas.0507626102

Kayhan, E., Meyer, M., O'Reilly, J. X., Hunnius, S., & Bekkering, H. (2019). Nine-month-old infants update their predictive models of a changing environment. *Developmental Cognitive Neuroscience*, *38*, 100680. https://doi.org/10.1016/J.DCN.2019.100680

Keller, E. (2000). Decoding the genetic program: Or, some cicular logic in the logic of circularity. In P. Beurton, R. Falk, & H-J. Rheinberger (Eds.), *The Concept of the Gene in Development and Evolution: Historical and Epistemological Perspectives* (pp. 159–177). Cambridge University Press. https://doi.org/10.1017/CBO9780511527296

King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R., & Montague, P. R. (2005). Getting to know you: Reputation and trust in a two-person economic exchange. *Science*, *308*(5718), 78-83. https://doi.org/10.1029/2003PA000992

Kingsbury, L., Huang, S., Wang, J., Gu, K., Golshani, P., Wu, Y. E., & Hong, W. (2019). Correlated neural activity and encoding of behavior across brains of socially interacting animals. *Cell*, *178*(2), 429-446.e16. https://doi.org/10.1016/j.cell.2019.05.022

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition, 83*(2), B35-B42. https://doi.org/10.1016/S0010-0277(02)00004-5

Klimesch, W., Schimke, H., Doppelmayr, M., Ripper, B., Schwaiger, J., & Pfurtscheller, G. (1996). Event-related desynchronization (ERD) and the Dm effect: Does alpha desynchronization during encoding predict later recall performance? *International Journal of Psychophysiology*, *24*(1–2), 47–60. https://doi.org/10.1016/S0167-8760(96)00054-2

Klimesch, W, Doppelmayr, M., Pachinger, T., & Russegger, H. (1997). Event-related desynchronization in the alpha band and the processing of semantic information. *Cognitive Brain Research*, *6*(2), 83–94. https://doi.org/10.1016/S0926-6410(97)00018-9

Klimesch, W, Doppelmayr, M., Russegger, H., & Pachinger, T. (1996). Theta band power in the human scalp EEG and the encoding of new information. *NeuroReport*, *7*(7), 1235–1240. https://doi.org/10.1097/00001756-199605170-00002

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

Koenig, M. A., Clément, F., & Harris, P. L. (2004). Trust in testimony: children's use of true and false statements. *Psychological Science*, *15*(10), 694–698. https://doi.org/10.1111/j.0956-7976.2004.00742.x

Koike, T., Sumiya, M., Nakagawa, E., Okazaki, S., & Sadato, N. (2019). What makes eye contact special? neural substrates of on-line mutual eye-gaze: A hyperscanning fMRI study. *ENeuro*, *6*(1). https://doi.org/10.1523/ENEURO.0284-18.2019

Koike, T., Tanabe, H. C., Adachi-Abe, S., Okazaki, S., Nakagawa, E., Sasaki, A. T., … Sadato, N. (2019). Role of the right anterior insular cortex in joint attention-related identification with a partner. *Social Cognitive and Affective Neuroscience*, *14*(10), 1131–1145. https://doi.org/10.1093/scan/nsz087

Kolb, B., Harker, A., & Gibb, R. (2017). Principles of plasticity in the developing brain. *Developmental Medicine & Child Neurology*, *59*(12), 1218–1223. https://doi.org/10.1111/dmcn.13546

Kopp, F., & Lindenberger, U. (2011). Effects of joint attention on long-term memory in 9-month-old infants: An event-related potentials study. *Developmental Science, 14(4), 660-672*. https://doi.org/10.1111/j.1467-7687.2010.01010.x

Köster, M., Haese, A., & Czernochowski, D. (2017). Neuronal oscillations reveal the processes underlying intentional compared to incidental learning in children and young adults. *PloS one*, *12*(8), e0182540. https://doi.org/10.1371/journal.pone.0182540.

Köster, M., Kayhan, E., Langeloh, M., & Hoehl, S. (2020). Making sense of the world: Infant learning from a predictive processing perspective. *Perspectives on Psychological Science, 15(3), 562-571*. https://doi.org/10.1177/1745691619895071

Köster, M., Langeloh, M., & Hoehl, S. (2019). Visually entrained theta oscillations increase for unexpected events in the infant brain. *Psychological Science*, *30*(11), 1656–1663. https://doi.org/10.1177/0956797619876260

Köster, M., Langeloh, M., Michel, C., & Hoehl, S. (2021). Young infants process prediction errors at the theta rhythm. *NeuroImage*, *236*, 118074. https://doi.org/10.1016/j.neuroimage.2021.118074

Kourtis, D., Sebanz, N., & Knoblich, G. (2013). Predictive representation of other people's actions in joint action planning: An EEG study. *Social Neuroscience*, *8*(1), 31–42. https://doi.org/10.1080/17470919.2012.694823

Kovács, Á. M., Téglás, E., & Endress, A. D. (2010). The social sense: susceptibility to others' beliefs in human infants and adults. *Science*, *330*(6012), 1830–1834. https://doi.org/10.1126/SCIENCE.1190792

Kuhl, P. K. (2007). Is speech learning 'gated' by the social brain? *Developmental Science*, *10*(1), 110–120. https://doi.org/10.1111/j.1467-7687.2007.00572.x

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1493), 979–1000. https://doi.org/10.1098/rstb.2007.2154

Kuhl, P. K., Tsao, F.M., & Liu, H. M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences*, *100*(15), 9096-9101. https://doi.org/10.1073/pnas.1532872100

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*(4427), 203-205. https://doi.org/10.1126/science.7350657

Kutsuki, A., Egami, S., Ogura, T., Nakagawa, K., Kuroki, M., & Itakura, S. (2007). Developmental changes of referential looks in 7-and 9-month-olds: A transition from dyadic to proto-referential looks. *Psychologia*, *50*(4), 319–329. https://doi.org/10.2117/psysoc.2007.319

Lachaux, J. P., Rodriguez, E., Martinerie, J., & Varela, F. J. (1999). Measuring phase synchrony in brain signals. *Hum Brain Mapping*, *8*, 194–208. https://doi.org/10.1002/(SICI)1097-0193(1999)8:4

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*, 863. https://doi.org/10.3389/FPSYG.2013.00863/ABSTRACT

Landers, W. F. (1971). Effects of differential experience on infants' performance in a Piagetian Stage IV object-concept task. *Developmental Psychology*, *5*(1), 48–54. https://doi.org/10.1037/h0031064

Langeloh, M., Buttelmann, D., Pauen, S., & Hoehl, S. (2020). 12- to 14-month-olds expect unconstrained agents to act efficiently: Event-related potential (ERP) evidence from the head-touch paradigm. *Developmental Psychology*, *56*(7), 1252–1267. https://doi.org/10.1037/DEV0000934

Leong, V., Byrne, E., Clackson, K., Georgieva, S., Lam, S., & Wass, S. (2017). Speaker gaze increases information coupling between infant and adult brains. *Proceedings of the National Academy of Sciences, 114*(50), 13290-13295. https://doi.org/10.1073/pnas.1702493114

Li, Y., Chen, R., Turel, O., Feng, T., Zhu, C. Z. Z., & He, Q. (2020). Dyad sex composition effect on inter-brain synchronization in face-to-face cooperation. *Brain Imaging and Behavior, 15*(3), 1667-1675. https://doi.org/10.1007/s11682-020-00361-z

Lillard, A. S., Lerner, M. D., Hopkins, E. J., Dore, R. A., Smith, E. D., & Palmquist, C. M. (2012). The impact of pretend play on children’s development: a review of the evidence. *Psychology Bulletin*, *139*(1), 1–34. https://doi.org/10.1037/a0029321

Lindenberger, U., Li, S. C. C., Gruber, W., & Müller, V. (2009). Brains swinging in concert: Cortical phase synchronization while playing guitar. *BMC Neuroscience*, *10*(1), 1-12). https://doi.org/10.1186/1471-2202-10-22

Lloyd-Fox, S., Blasi, A., McCann, S., Rozhko, M., Katus, L., Mason, L., … Elwell, C. E. (2019). Habituation and novelty detection fNIRS brain responses in 5- and 8-month-old infants: The Gambia and UK. *Developmental Science*, *22*(5), e12817. https://doi.org/10.1111/desc.12817

Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature*, *453*(7197), 869–878. https://doi.org/10.1038/nature06976

Longo, M. R., & Bertenthal, B. I. (2006). Common coding of observation and execution of action in 9-month-old infants. *Infancy*, *10*(1), 43–59. https://doi.org/10.1207/s15327078in1001_3

Loucks, J., & Sommerville, J. A. (2012). The role of motor experience in understanding action function: The case of the precision grasp. *Child Development*, *83*(3), 801–809. https://doi.org/10.1111/j.1467-8624.2012.01735.x

Luchies, L. B., Rusbult, C. E., Eastwick, P. W., Wieselquist, J., Kumashiro, M., Coolsen, M. K., & Finkel, E. J. (2013). Trust and biased memory of transgressions in romantic relationships. *Journal of Personality and Social Psychology, 104*(4), 673. https://doi.org/10.1037/a0031054

Luck, S. J. (2005). *An introduction to the event-related potential technique*. MIT Press.

Luo, Y., & Baillargeon, R. (2005). When the ordinary seems unexpected: Evidence for incremental physical knowledge in young infants. *Cognition, 95*(3), 297-328. https://doi.org/10.1016/j.cognition.2004.01.010

Maccotta, L., & Buckner, R. L. (2004). Evidence for neural effects of repetition that directly correlate with behavioral priming. *Journal of Cognitive Neuroscience*, *16*(9), 1625–1632. https://doi.org/10.1162/0898929042568451

Makowski, D., Ben-Shachar, M., & Lüdecke, D. (2019). bayestestR: Describing effects and their uncertainty, existence and significance within the Bayesian framework. *Journal of Open Source Software*, *4*(40), 1541. https://doi.org/10.21105/joss.01541

Malcuit, G., Pomerleau, A., & Lamarre, G. (1988). Habituation, visual fixation and cognitive activity in infants: A critical analysis and attempt at a new formulation. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, *8*(5), 415–440.

Marcovitch, S., & Zelazo, P. D. (1999). The A-not-B error: Results from a logistic meta-analysis. *Child Development*, *70*(6), 1297–1313. https://doi.org/10.1111/1467-8624.00095

Markova, G., Nguyen, T., & Hoehl, S. (2019). Neurobehavioral interpersonal synchrony in early development: The role of interactional rhythms. *Frontiers in Psychology*, *10*, 2078. https://doi.org/10.3389/fpsyg.2019.02078

Maye, J., Werker, J. F., & Gerken, L. A. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*(3). https://doi.org/10.1016/S0010-0277(01)00157-3

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*(3), 419–457. https://doi.org/10.1037/0033-295X.102.3.419

McNally, R., & Alborz, A. (2004). Developing methods for systematic reviewing in health services delivery and organization: an example from a review of access to health care for people with learning disabilities. Part 1. Identifying the literature. *Health Information and Libraries Journal*, *21*(3), 182–02. https://doi.org/10.1111/j.1471-1842.2004.00512.x

Meadows, S. (2009). *The child as social person*. Routledge. https://doi.org/10.4324/9780203861059

Meltzoff, A. N. (2007). "Like me": A foundation for social cognition. *Developmental Science*, *10*(1), 126-134. https://doi.org/10.1111/j.1467-7687.2007.00574.x

Meltzoff, A. N., & Brooks, R. (2007). Eyes wide shut: The importance of eyes in infant gaze following and understanding other minds. In R. Flom, K. Lee, & D. Muir (Eds.), *Gaze following: Its development and significance* (pp. 217–241). Erlbaum.

Michael, J. (2011). Interactionism and mindreading. *Review of Philosophy and Psychology*, *2*(3), 559–578. https://doi.org/10.1007/s13164-011-0066-z

Michel, C., Kaduk, K., & Ní Choisdealbha, Á. (2017). Event-related potentials discriminate familiar and unusual goal outcomes in 5-month-olds and adults. *Developmental Psychology*, *53*(10), 1833. https://doi.org/10.1037/dev0000376

Michel, C., Stets, M., Parise, E., Reid, V. M., Striano, T., & Hoehl, S. (2015). Theta- and alpha-band EEG activity in response to eye gaze cues in early infancy. *NeuroImage*, *118*, 576–583. https://doi.org/10.1016/j.neuroimage.2015.06.042

Mills, C. M. (2013). Knowing when to doubt: developing a critical stance when learning from others. *Developmental Psychology, 49*(3), 404. https://doi.org/10.1037/a0029500

Molden, D. C., & Finkel, E. J. (2010). Motivations for promotion and prevention and the role of trust and commitment in interpersonal forgiveness. *Journal of Experimental Social Psychology*, *46*(2), 255-268. https://doi.org/10.1016/j.jesp.2009.10.014

Moll, H., Carpenter, M., & Tomasello, M. (2007). Fourteen-month-olds know what others experience only in joint engagement. *Developmental Science*, *10*(6), 826–835. https://doi.org/10.1111/j.1467-7687.2007.00615.x

Moll, H., & Tomasello, M. (2004). 12- and 18-month-old infants follow gaze to spaces behind barriers. *Developmental Science*, *7*(1). https://doi.org/10.1111/j.1467-7687.2004.00315.x

Moll, H., & Tomasello, M. (2007). How 14- and 18-month-olds know what others have experienced. *Developmental Psychology*, *43*(2), 309–317. https://doi.org/10.1037/0012-1649.43.2.309

Monroy, C., Gerson, S., & Hunnius, S. (2017). Infants’ motor proficiency and statistical learning for actions. *Frontiers in Psychology*, *8*, 2174. https://doi.org/10.3389/fpsyg.2017.02174

Montague, P. R., Berns, G. S., Cohen, J. D., McClure, S. M., Pagnoni, G., Dhamala, M., … Fisher, R. E. (2002). Hyperscanning: Simultaneous fMRI during linked social interactions. *NeuroImage, 16*(4), 1159-1164. https://doi.org/10.1006/nimg.2002.1150

Mullen, B. (2013). *Advanced basic meta-analysis: Version 1.10*. Psychology Press.

Müller, M. M., Picton, T. W., Valdes-Sosa, P., Riera, J., Teder-Sälejärvi, W. A., & Hillyard, S. A. (1998). Effects of spatial selective attention on the steady-state visual evoked potential in the 20–28 Hz range. *Cognitive Brain Research*, *6*(4), 249–261. https://doi.org/10.1016/S0926-6410(97)00036-0

Munakata, Y. (1997). Perseverative reaching in infancy: The roles of hidden toys and motor history in the AB task. *Infant Behavior and Development*, *20*(3), 405–416. https://doi.org/10.1016/S0163-6383(97)90011-4

Munakata, Y. (1998). Infant perseveration and implications for object permanence theories: A PDP model of the AB task. *Developmental Science*, *1*(2), 161–184. https://doi.org/10.1111/1467-7687.00021

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

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

Mundy, P., & Gomes, A. (1998). Individual differences in joint attention skill development in the second year. *Infant Behavior and Development*, *21*(3), 469–482. https://doi.org/10.1016/S0163-6383(98)90020-0

Murray, L., & Trevarthen, C. (1985). Emotional regulation of interactions between two-month-olds and their mothers. In T. Field & N. A. Fox (Eds.), *Social perception in infants* (pp. 177–197). Ablex.

Nagy, W. E., & Scott, J. A. (2000). Vocabulary processes. In M. Kamil, P. Mosenthal, P. Pearson, & R. Barr (Eds.), *Handbook of Reading Research* Vol. 3 (pp. 269–284). Routledge. https://doi.org/10.4324/9781315200613

Nakano, T., Watanabe, H., Homae, F., & Taga, G. (2009). Prefrontal cortical involvement in young infants’ analysis of novelty. *Cerebral Cortex*, *19*(2), 455–463. https://doi.org/10.1093/cercor/bhn096

Nazzi, T., & Bertoncini, J. (2003). Before and after the vocabulary spurt: Two modes of word acquisition? *Developmental Science*, *6*(2), 136–142. https://doi.org/10.1111/1467-7687.00263

Neale, D., Clackson, K., Georgieva, S., Dedetas, H., Scarpate, M., Wass, S., & Leong, V. (2018). Toward a neuroscientific understanding of play: A dimensional coding framework for analyzing infant-adult play patterns. *Frontiers in Psychology*, *9*, 273. https://doi.org/10.3389/fpsyg.2018.00273

Neisser, U. (1980). On “social knowing.” *Personality and Social Psychology Bulletin*, *6*(4), 601–605. https://doi.org/10.1177/014616728064012

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

Nelson, C. A. (1997). Electrophysiological correlates of memory development in the first year of life. In H. W. Reese & M. D. Franzen (Eds.), *Biological and neuropsychological mechanisms* (pp. 95-131). Mahwah, NJ: Erlbaum.

Nelson, C. A. (1997). Electrophysiological correlates of memory development in the first year of life. In H. M. Reese & M. D. Franzen (Eds.), *Biological and neuropsychological mechanisms: Life-span developmental psychology* (pp. 95–131). Lawrence Erlbaum Associates, Inc.

Nelson, C. A. (1995). The ontogeny of human memory: A cognitive neuroscience perspective. *Developmental Psychology*, *31*(5), 723–738. https://doi.org/10.1037/0012-1649.31.5.723

Nelson, C. A., & Collins, P. F. (1992). Neural and behavioral correlates of visual recognition memory in 4- and 8-month-old infants. *Brain and Cognition, 19*(1), 105-121. https://doi.org/10.1016/0278-2626(92)90039-O

Nelson, C. A., Furtado, E. A., Fox, N. A., & Zeanah, C. H. (2009). The deprived human brain. *American Scientist*, *97*(3), 222–229. https://doi.org/10.1511/2009.78.222

Nelson, C. A., & Collins, P. F. (1991). Event-related potential and looking-time analysis of infants’ responses to familiar and novel events: Implications for visual recognition memory. *Developmental Psychology, 27*(1), 50. https://doi.org/10.1037/0012-1649.27.1.50

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

Neuper, C., & Pfurtscheller, G. (2001). Event-related dynamics of cortical rhythms: Frequency-specific features and functional correlates. *International Journal of Psychophysiology*, *43*, 41–58. https://doi.org/10.1016/S0167-8760(01)00178-7

Nguyen, T., Bánki, A., Markova, G., & Hoehl, S. (2020). Studying parent-child interaction with hyperscanning. *Progress in Brain Research, 254*, 1-24. https://doi.org/10.1016/bs.pbr.2020.05.003

Ní Choisdealbha, Á., & Reid, V. M. (2014). The developmental cognitive neuroscience of action: semantics, motor resonance and social processing. *Experimental Brain Research, 232*(6), 1585–1597. https://doi.org/10.1007/S00221-014-3924-Y

Nikkel, L., & Karrer, R. (1994). Differential effects of experience on the ERP and behavior of 6-month-old infants: trends during repeated stimulus presentations. *Developmental Neuropsychology*, *10*(1), 1-11. https://doi.org/10.1080/87565649409540561

Noreika, V., Georgieva, S., Wass, S., & Leong, V. (2020). 14 challenges and their solutions for conducting social neuroscience and longitudinal EEG research with infants. *Infant Behavior and Development*, *58*. https://doi.org/10.1016/j.infbeh.2019.101393

Norman, K. A., & O’Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. *Psychological Review*, *110*(4), 611–646. https://doi.org/10.1037/0033-295X.110.4.611

Nozawa, T., Sasaki, Y., Sakaki, K., Yokoyama, R., & Kawashima, R. (2016). Interpersonal frontopolar neural synchronization in group communication: An exploration toward fNIRS hyperscanning of natural interactions. *NeuroImage*, *133*, 484–497. https://doi.org/10.1016/j.neuroimage.2016.03.059

Nunez, P. L., & Srinivasan, R. (2006). *Electric fields of the brain: The neurophysics of EEG*. (2nd edition). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195050387.001.0001

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

O’Reilly, R. C., & Norman, K. A. (2002). Hippocampal and neocortical contributions to memory: Advances in the complementary learning systems framework. *Trends in Cognitive Sciences*, *6*(12), 505–510. https://doi.org/10.1016/S1364-6613(02)02005-3

Onishi, K. H., Baillargeon, R., & Leslie, A. M. (2007). 15-month-old infants detect violations in pretend scenarios. *Acta Psychologica*, *124*(1), 106–128. https://doi.org/10.1016/j.actpsy.2006.09.009

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., & Elam, M. (2006). EEG theta rhythm in infants and preschool children. *Clinical Neurophysiology*, *117*(5), 1047–1062. https://doi.org/10.1016/j.clinph.2005.12.027

Orekhova, 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

Osaka, N., Minamoto, T., Yaoi, K., Azuma, M., Shimada, Y. M., & Osaka, M. (2015). How two brains make one synchronized mind in the inferior frontal cortex: fNIRS-based hyperscanning during cooperative singing. *Frontiers in Psychology*, *6*, 1811. https://doi.org/10.3389/fpsyg.2015.01811

Ozer, O., & Zheng, Y. (2019). Trust and trustworthiness. In K. Donohue, E. Katok, & S. Leider (Eds.), *The handbook of behavioral operations.* (pp. 489–523). Wiley. https://doi.org/10.1002/9781119138341

Pace, A., Carver, L. J., & Friend, M. (2013). Event-related potentials to intact and disrupted actions in children and adults. *Journal of Experimental Child Psychology*, *116*(2), 453–470. https://doi.org/10.1016/j.jecp.2012.10.013

Pallant, J. (2010). *SPSS Survival manual: A step by step guide to data analysis using SPSS* (4th Edition). McGraw-Hill International.

Pan, Y., Cheng, X., Zhang, Z., Li, X., & Hu, Y. (2017). Cooperation in lovers: An fNIRS-based hyperscanning study. *Human Brain Mapping*, *38*(2), 831–841. https://doi.org/10.1002/hbm.23421

Papaioannou, D., Sutton, A., Carroll, C., Booth, A., & Wong, R. (2010). Literature searching for social science systematic reviews: Consideration of a range of search techniques. *Health Information and Libraries Journal*, *27*(2), 114–122. https://doi.org/10.1111/j.1471-1842.2009.00863.x

Parise, E., & Csibra, G. (2012). Electrophysiological evidence for the understanding of maternal speech by 9-month-old infants. *Psychological Science*, *23*(7), 728–733. https://doi.org/10.1177/0956797612438734

Parise, E., Reid, V. M., Stets, M., & Striano, T. (2008). Direct eye contact influences the neural processing of objects in 5-month-old infants. *Social Neuroscience*, *3*(2), 141–150. https://doi.org/10.1080/17470910701865458

Patel, C. J., Burford, B., & Ioannidis, J. P. A. (2015). Assessment of vibration of effects due to model specification can demonstrate the instability of observational associations. *Journal of Clinical Epidemiology*, *68*(9), 1046–1058. https://doi.org/10.1016/j.jclinepi.2015.05.029

Paulus, M., Hunnius, S., Vissers, M., & Bekkering, H. (2011). Imitation in infancy: Rational or motor resonance? *Child Development*, *82*(4), 1047–1057. https://doi.org/10.1111/j.1467-8624.2011.01610.x

Perone, S., & Gartstein, M. A. (2019). Relations between dynamics of parent-infant interactions and baseline EEG functional connectivity. *Infant Behavior and Development*, *57*, 101344. https://doi.org/10.1016/j.infbeh.2019.101344

Piaget, J. (1954). *The construction of reality in the child*. Basic Books. https://doi.org/10.1037/11168-000

Piaget, J. (1966). *Psychology of intelligence*. Routledge.

Pinti, P., Tachtsidis, I., Hamilton, A., Hirsch, J., Aichelburg, C., Gilbert, S., & Burgess, P. W. (2018). The present and future use of functional near-infrared spectroscopy (fNIRS) for cognitive neuroscience. *Annals of the New York Academy of Sciences*, *1464*(1), 5–29. https://doi.org/10.1111/nyas.13948

Posner, M. I. (2004). *Cognitive neuroscience of attention*. The Guilford Press.

Poulin-Dubois, D., Brooker, I., & Polonia, A. (2011). Infants prefer to imitate a reliable person. *Infant Behavior and Development*, *34*(2), 303–309. https://doi.org/10.1016/j.infbeh.2011.01.006

R Core Team. (2014). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*. Vienna, Austria. https://www.R-project.org/

Redcay, E., & Schilbach, L. (2019). Using second-person neuroscience to elucidate the mechanisms of social interaction. *Nature Reviews Neuroscience*, *20*(8), 495–505. https://doi.org/10.1038/s41583-019-0179-4

Reddy, V. (2008). *How infants know minds*. Harvard University Press.

Reid, V. M., Striano, T., & Iacoboni, M. (2011). Neural correlates of dyadic interaction during infancy. *Developmental Cognitive Neuroscience*, *1*(2), 124–130. https://doi.org/10.1016/j.dcn.2011.01.001

Reid, V. M., Csibra, G., Belsky, J., & Johnson, M. H. (2007). Neural correlates of the perception of goal-directed action in infants. *Acta Psychologica, 124*(1), 129-138. https://doi.org/10.1016/j.actpsy.2006.09.010

Reid, V. M., Hoehl, S., Grigutsch, M., Groendahl, A., Parise, E., & Striano, T. (2009). The neural correlates of infant and adult goal prediction: Evidence for semantic processing systems. *Developmental Psychology, 45*(3), 620. https://doi.org/10.1037/a0015209

Reid, V. M., & Striano, T. (2005). Adult gaze influences infant attention and object processing: Implications for cognitive neuroscience. *European Journal of Neuroscience*, *21*(6), 1763–1766. https://doi.org/10.1111/j.1460-9568.2005.03986.x

Reid, V. M., & Striano, T. (2007). The directed attention model of infant social cognition. *European Journal of Developmental Psychology*, *4*(1), 100–110. https://doi.org/10.1080/17405620601005648

Reid, V. M., & Striano, T. (2008). N400 involvement in the processing of action sequences. *Neuroscience Letters*, *433*(2), 93–97. https://doi.org/10.1016/j.neulet.2007.12.066

Reid, V. M., Striano, T., Kaufman, J., & Johnson, M. H. (2004). Eye gaze cuing facilitates neural processing of objects in 4 month old infants. *NeuroReport*, *15*(16), 2553–2555. https://doi.org/10.1097/00001756-200411150-00025

Reindl, V., Gerloff, C., Scharke, W., & Konrad, K. (2018). Brain-to-brain synchrony in parent-child dyads and the relationship with emotion regulation revealed by fNIRS-based hyperscanning. *NeuroImage*, *178*, 493–502. https://doi.org/10.1016/j.neuroimage.2018.05.060

Rempel, J. K., Holmes, J. G., & Zanna, M. P. (1985). Trust in close relationships. *Journal of Personality and Social Psychology, 49*(1), 95. https://doi.org/10.1037/0022-3514.49.1.95

Reynolds, G. D., Courage, M. L., & Richards, J. E. (2010). Infant attention and visual preferences: converging evidence from behavior, event-related potentials, and cortical source localization. *Developmental psychology*, *46*(4), 886–904. https://doi.org/10.1037/a0019670

Reynolds, G. D., & Richards, J. E. (2005). Familiarization, attention, and recognition memory in infancy: An event-related potential and cortical source localization study. *Developmental Psychology*, *41*(4), 598–615. https://doi.org/10.1037/0012-1649.41.4.598

Richards, J. E. (2003). Attention affects the recognition of briefly presented visual stimuli in infants: An ERP study. *Developmental Science, 6*(3), 312-328. https://doi.org/10.1111/1467-7687.00287

Richardson, R. C. (1979). Functionalism and reductionism. *Philosophy of Science*, *46(4)*, 533–558. Retrieved from https://www.jstor.org/stable/187248

Rivera-Gaxiola, M., Silva-Pereyra, J., & Kuhl, P. K. (2005). Brain potentials to native and non-native speech contrasts in 7- and 11-month-old American infants. *Developmental Science*, *8*(2), 162–172. https://doi.org/10.1111/j.1467-7687.2005.00403.x

Rochat, P. (2001). *The infant's world*. Harvard University Press.

Rochat, P. (2014). *Early Social Cognition Understanding Others in the First Months of Life.* Psychology Press.

Rugg, M. D., & Coles, M. G. H. (1996). *Electrophysiology of Mind: Event-Related Brain Potentials and Cognition*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198524168.001.0001

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

Saby, J. N., Meltzoff, A. N., & Marshall, P. J. (2013). Infants' somatotopic neural responses to seeing human actions: I've got you under my skin. *PLoS ONE*, *8*(10), e77905. https://doi.org/10.1371/journal.pone.0077905

Saffran, J. R. (2001). Words in a sea of sounds: the output of infant statistical learning. *Cognition*, *81*, 149–169. https://doi.org/10.1016/S0010-0277(01)00132-9

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928. https://doi.org/10.1126/science.274.5294.1926

Saffran, J. R., & Kirkham, N. Z. (2018). Infant statistical learning. *Annual Review of Psychology*, *69*(1), 181–203. https://doi.org/10.1146/annurev-psych-122216-011805

Saito, D. N., Tanabe, H. C., Izuma, K., Hayashi, M. J., Morito, Y., Komeda, H., … Sadato, N. (2010). “Stay tuned”: Inter-individual neural synchronization during mutual gaze and joint attention. *Frontiers in Integrative Neuroscience*, *4*, 1–12. https://doi.org/10.3389/fnint.2010.00127

Sameroff, A. J. (2009). *The transactional model of development: how children and contexts shape each other*. American Psychological Association. https://doi.org/10.1037/11877-001

Santamaria, L., Noreika, V., Georgieva, S., Clackson, K., Wass, S., & Leong, V. (2020). Emotional valence modulates the topology of the parent-infant inter-brain network. *NeuroImage*, *207*. https://doi.org/10.1016/j.neuroimage.2019.116341

Sarkart, S. (2000). Information in genetics and developmental biology: Comments on Maynard Smith. *Philosophy of Science*, *67*, 208–213. Retrieved from https://www.jstor.org/stable/188720

Saville, D. J. (1990). Multiple comparison procedures: the practical solution. *The American Statistician*, *44*(2), 174-180. https://doi.org/10.2307/2684163

Sawyer, R. K. (2002). Emergence in psychology: Lessons from the history of non-reductionist science. *Human Development*, *45*(1), 2–28. https://doi.org/10.1159/000048148

Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature*, *253*(5489), 265–266. https://doi.org/10.1038/253265a0

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences*, *36*, 393–414. https://doi.org/10.1017/S0140525X12000660

Schneiderman, I., Zagoory-Sharon, O., Leckman, J. F., & Feldman, R. (2012). Oxytocin during the initial stages of romantic attachment: Relations to couples’ interactive reciprocity. *Psychoneuroendocrinology*, *37*(8), 1277–1285. https://doi.org/10.1016/j.psyneuen.2011.12.021

Scholkmann, F., Holper, L., Wolf, U., & Wolf, M. (2013). A new methodical approach in neuroscience: Assessing inter-personal brain coupling using functional near-infrared imaging (fNIRI) hyperscanning. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00813

Schöner, G., & Thelen, E. (2006). Using dynamic field theory to rethink infant habituation. *Psychological Review*, *113*(2), 273–299. https://doi.org/10.1037/0033-295X.113.2.273

Schubotz, R. I. (2015). Prediction and expectation. In A. W. Toga (Ed.), *Brain Mapping: An Encyclopedic Reference Volume 3* (pp. 295–302). Academic Press.

Schütte, F., Mani, N., & Behne, T. (2020). Retrospective inferences in selective trust. *Royal Society Open Science*, *7*(2), 191451. https://doi.org/10.1098/rsos.191451

Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, *18*(9), 668–671. https://doi.org/10.1016/J.CUB.2008.03.059

Senju, A., Johnson, M. H., & Csibra, G. (2006). The development and neural basis of referential gaze perception. *Social Neuroscience*, *1*(3–4), 220–234. https://doi.org/10.1080/17470910600989797

Shaw, D. J., Czekóová, K., Staněk, R., Mareček, R., Urbánek, T., Špalek, J., … Brázdil, M. (2018). A dual-fMRI investigation of the iterated Ultimatum Game reveals that reciprocal behaviour is associated with neural alignment. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-29233-9

Shimada, S., & Hiraki, K. (2006). Infant's brain responses to live and televised action. *NeuroImage*, *32*(2), 930–939. https://doi.org/10.1016/j.neuroimage.2006.03.044

Shneidman, L. A., Buresh, J. S., Shimpi, P. M., Knight-Schwarz, J., & Woodward, A. L. (2009). Social experience, social attention and word learning in an overhearing paradigm. *Language Learning and Development*, *5*(4), 266–281. https://doi.org/10.1080/15475440903001115

Siemens, W., Meerpohl, J. J., Rohe, M. S., Buroh, S., Schwarzer, G., & Becker, G. (2021). Reevaluation of statistically significant meta-analyses in advanced cancer patients using the Hartung-Knapp method and prediction intervals–a methodological study. *Research Synthesis Methods*. https://doi.org/10.1002/jrsm.1543

Silberzahn, R., Uhlmann, E. L., Martin, D. P., Anselmi, P., Aust, F., Awtrey, E., … Nosek, B. A. (2018). Many analysts, one data set: Making transparent how variations in analytic choices affect results. *Advances in Methods and Practices in Psychological Science*, *1*(3), 337–356. https://doi.org/10.1177/2515245917747646

Simonsohn, U., Simmons, J. P., & Nelson, L. D. (2015). Specification curve: Descriptive and inferential statistics on all reasonable specifications. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2694998

Singer, W. (1995). Development and plasticity of cortical processing architectures. *Science*, *270*(5237), 758–764. https://doi.org/10.1126/SCIENCE.270.5237.758

Siposova, B., & Carpenter, M. (2019). A new look at joint attention and common knowledge. *Cognition*, *189*, 260–274. https://doi.org/10.1016/j.cognition.2019.03.019

Sirois, S., & Mareschal, D. (2002). Models of habituation in infancy. *Trends in Cognitive Sciences*, *6*(7), 293–298. https://doi.org/10.1016/S1364-6613(02)01926-5

Smith, E. S., Elliott, D., Killick, R., Crawford, T. J., Kidby, S., & Reid, V. M. (2021). Infants oscillatory frequencies change during free-play. *Infant Behavior and Development*, *64*, 101612. https://doi.org/10.1016/j.infbeh.2021.101612

Smith, L. B. (2005). Cognition as a dynamic system: Principles from embodiment. *Developmental Review*, *25*(3–4), 278–298. https://doi.org/10.1016/j.dr.2005.11.001

Smith, L. B., & Thelen, E. (1994). *A dynamic systems approach to the development of cognition and action*. MIT Press.

Smith, L. B., Thelen, E., Titzer, R., & McLin, D. (1999). Knowing in the context of acting: The task dynamics of the A-not-B error. *Psychological Review*, *106*(2), 235–260. https://doi.org/10.1037/0033-295X.106.2.235

Smith, L. B., & Yu, C. (2012). Visual attention is not enough: Individual differences in statistical word-referent learning in infants. *Language Learning and Development, 9*(1), 25-49. https://doi.org/10.1080/15475441.2012.707104

Snyder, K., Webb, S. J., & Nelson, C. A. (2002). Theoretical and methodological implications of variability in infant brain response during a recognition memory paradigm. *Infant Behavior and Development, 25*(4), 466-494. https://doi.org/10.1016/S0163-6383(02)00146-7

Sommerville, J. A., Woodward, A. L., & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition*, *96*(1), 1–11. https://doi.org/10.1016/j.cognition.2004.07.004

Southgate, V., Csibra, G., Kaufman, J., & Johnson, M. H. (2008). Distinct processing of objects and faces in the infant brain. *Journal of Cognitive Neuroscience*, *20*(4), 741–749. https://doi.org/10.1162/jocn.2008.20052

Southgate, V., Johnson, M. H., Karoui, I. El, & Csibra, G. (2010). Motor system activation reveals infants' on-line prediction of others' goals. *Psychological Science*, *21*(3), 355–359. https://doi.org/10.1177/0956797610362058

Spelke, E. S. (1994). Initial knowledge: six suggestions. *Cognition*, *50*(1–3), 431–445. https://doi.org/10.1016/0010-0277(94)90039-6

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review*, *99*(4), 605–632. https://doi.org/10.1037/0033-295X.99.4.605

Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science*, *10*(1), 89–96. https://doi.org/10.1111/j.1467-7687.2007.00569.x

Spiegelhalder, K., Ohlendorf, S., Regen, W., Feige, B., van Elst, L. T., Weiller, C., … Tüscher, O. (2014). Interindividual synchronization of brain activity during live verbal communication. *Behavioural Brain Research*, *258*, 75–79. https://doi.org/10.1016/j.bbr.2013.10.015

St. John, A. M., Kao, K., Choksi, M., Liederman, J., Grieve, P. G., & Tarullo, A. R. (2016). Variation in infant EEG power across social and nonsocial contexts. *Journal of Experimental Child Psychology*, *152*. https://doi.org/10.1016/j.jecp.2016.04.007

Stahl, A. E., & Feigenson, L. (2015). Observing the unexpected enhances infants' learning and exploration. *Science*, *348*(6230), 91–94. https://doi.org/10.1126/science.aaa3799

Stahl, A. E., & Feigenson, L. (2019). Violations of core knowledge shape early learning. *Topics in Cognitive Science, 11*(1), 136-153. https://doi.org/10.1111/tops.12389

Steegen, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016). Increasing transparency through a multiverse analysis. *Perspectives on Psychological Science*, *11*(5), 702–712. https://doi.org/10.1177/1745691616658637

Stenberg, G. (2009). Selectivity in infant social referencing. *Infancy*, *14*(4), 457–473. https://doi.org/10.1080/15250000902994115

Stenberg, G. (2013). Do 12-month-old infants trust a competent adult? *Infancy*, *18*(5), 873–904. https://doi.org/10.1111/infa.12011

Stenberg, G. (2017). Does contingency in adults' responding influence 12-month-old infants' social referencing? *Infant Behavior and Development*, *46*, 67-79. https://doi.org/10.1016/j.infbeh.2016.11.013

Stephens, G. J., Silbert, L. J., & Hasson, U. (2010). Speaker-listener neural coupling underlies successful communication. *Proceedings of the National Academy of Sciences, 107*(32), 14425-14430. https://doi.org/10.1073/pnas.1008662107

Stern, D. N. (1985). *The interpersonal world of the infant: a view from psychoanalysis and developmental psychology*. Basic Books.

Stets, M., Burt, M., & Reid, V. M. (2013). Infants need more variety - increased data acquisition with reduced participant attrition in infant ERP studies. *Frontiers in Psychology*, *4*, 117. https://doi.org/10.3389/fpsyg.2013.00117

Stets, M., & Reid, V. M. (2011). Infant ERP amplitudes change over the course of an experimental session: Implications for cognitive processes and methodology. *Brain and Development*, *33*(7), 558–568. https://doi.org/10.1016/J.BRAINDEV.2010.10.008

Stiles, J. (2017). Principles of brain development. *Cognitive Science*, *8*, 1–5. https://doi.org/10.1002/WCS.1402

Striano, T., Reid, V. M., & Hoehl, S. (2006). Neural mechanisms of joint attention in infancy. *European Journal of Neuroscience, 23*(10), 2819-2823. https://doi.org/10.1111/j.1460-9568.2006.04822.x

Stroganova, T. A., 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

Tanabe, H. C., Kosaka, H., Saito, D. N., Koike, T., Hayashi, M. J., Izuma, K., … Sadato, N. (2012). Hard to “tune in”: Neural mechanisms of live face-to-face interaction with high-functioning autistic spectrum disorder. *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00268

Tass, P., Rosenblum, M. G., Weule, J., Kurths, J., Pikovsky, A., Volkmann, J., … Freund, H. J. (1998). Detection of n:m phase locking from noisy data: Application to magnetoencephalography. *Physical Review Letters*, *81*(15), 3291–3294. https://doi.org/10.1103/PhysRevLett.81.3291

Thatcher, R. W. (2012). Coherence, phase differences, phase shift, and phase lock in EEG/ERP analyses. *Developmental Neuropsychology*, *37*(6), 476–496. https://doi.org/10.1080/87565641.2011.619241

Thelen, E., Schöner, G., Scheier, C., & Smith, L. B. (2001). The dynamics of embodiment: A field theory of infant perseverative reaching. *Behavioral and Brain Sciences*, *24*(1), 1–34. https://doi.org/10.1017/S0140525X01003910

Thompson, E. (2007). *Mind in life: Biology, phenomenology, and the sciences of mind.* Harvard University Press.

Tincoff, R., & Jusczyk, P. W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, *10*(2), 172–175. https://doi.org/10.1111/1467-9280.00127

Tomasello, M. (2014). *A natural history of human thinking*. Harvard University Press.

Tomasello, M., & Haberl, K. (2003). Understanding attention: 12- and 18-month-olds know what is new for other persons. *Developmental Psychology*, *39*(5), 906–912. https://doi.org/10.1037/0012-1649.39.5.906

Tomasello, M., Kruger, A. C., & Ratner, H. H. (1993). Cultural learning. *Behavioral and Brain Sciences*, *16*(3), 495–511. https://doi.org/10.1017/S0140525X0003123X

Tomasello, M. (1999). *The cultural origins of human cognition*. Harvard University Press.

Topál, J., Gergely, G., Miklósi, Á., Erdohegyi, Á., & Csibra, G. (2008). Infants' perseverative search errors are induced by pragmatic misinterpretation. *Science*, *321*(5897), 1831–1834. https://doi.org/10.1126/science.1161437

Topor, M., Pickering, J., Mendes, A. B., Bishop, D., Büttner, F. C., Elsherif, M., … Nitschke, F. (2020). Non-Interventional, Reproducible, and Open (NIRO) Systematic Review guidelines. https://doi.org/10.17605/OSF.IO/F3BRW

Torrence, C., & Compo, G. P. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society*, *79*(1), 61–78. https://doi.org/10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2

Tronick, E. Z., Als, H., & Adamson, L. (1979). Mother-infant face-to-face communicative interaction. In M. Bullowa (Ed.), *Before Speech* (pp. 349–372). Cambridge University Press.

Tronick, E. Z. (1989). Emotions and emotional communication in infants. *American Psychologist*, *44*(2), 112–119. https://doi.org/10.1037/0003-066X.44.2.112

Tronick, E. Z., Heidelise, A., & Brazelton, T. B. (1980). Monadic phases: A structural descriptive analysis of infant-mother face to face interaction. *Merrill Palmer Quarterly, 26*(1), 3-24. https://doi.org/10.2307/23084074

Tummeltshammer, K. S., Wu, R., Sobel, D. M., & Kirkham, N. Z. (2014). Infants track the reliability of potential informants. *Psychological Science*, *25*(9), 1730–1738. https://doi.org/10.1177/0956797614540178

Turk-Browne, N. B., Yi, D. J., & Chun, M. M. (2006). Linking implicit and explicit memory: Common encoding factors and shared representations. *Neuron*, *49*(6), 917–927. https://doi.org/10.1016/j.neuron.2006.01.030

van Elk, M., van Schie, H. T., Hunnius, S., Vesper, C., & Bekkering, H. (2008). You'll never crawl alone: Neurophysiological evidence for experience-dependent motor resonance in infancy. *NeuroImage*, *43*(4), 808–814.

van Ijzendoorn, M. H., & Sagi-Schwartz, A. (2008). Cross-cultural patterns of attachment: Universal and contextual dimensions. In J. Cassidy & P. R. Shaver (Eds.), *Handbook of attachment: Theory, research, and clinical applications* (pp. 880–905). The Guilford Press.

Van Turennout, M., Ellmore, T., & Martin, A. (2000). Long-lasting cortical plasticity in the object naming system. *Nature Neuroscience*, *3*(12), 1329–1334. https://doi.org/10.1038/81873

Varela, F. J. (1997). Patterns of life: Intertwining identity and cognition. *Brain and Cognition*, *34*(1), 72–87. https://doi.org/10.1006/BRCG.1997.0907

Vehtari, A., Gabry, J., Magnusson, M., Yao, Y., Bürkner, P., Paananen, T., & Gelman, A. (2020). loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models. R package version 2.4.1.

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*, 1413–1432. https://doi.org/10.1007/s11222-016-9696-4

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Vogeley, K. (2017). Two social brains: Neural mechanisms of intersubjectivity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1727). https://doi.org/10.1098/rstb.2016.0245

Wahl, S., Michel, C., Pauen, S., & Hoehl, S. (2013). Head and eye movements affect object processing in 4-month-old infants more than an artificial orientation cue. *British Journal of Developmental Psychology*, *31*(2), 212–230. https://doi.org/10.1111/bjdp.12001

Walden, T. A., Kim, G., McCoy, C., & Karrass, J. (2007). Do you believe in magic? Infant's social looking during violations of expectations. *Developmental Science*, *10*(5), 654–663. https://doi.org/10.1111/j.1467-7687.2007.00607.x

Walden, T. A., & Ogan, T. A. (1988). The Development of social referencing. *Child Development*, *59*(5), 1230–1240. https://doi.org/10.2307/1130486

Wang, S. H., Baillargeon, R., & Paterson, S. (2005). Detecting continuity violations in infancy: A new account and new evidence from covering and tube events. *Cognition*, *95*(2), 129–173. https://doi.org/10.1016/j.cognition.2002.11.001

Wass, S. V., Clackson, K., Georgieva, S. D., Brightman, L., Nutbrown, R., & Leong, V. (2018). Infants' visual sustained attention is higher during joint play than solo play: Is this due to increased endogenous attention control or exogenous stimulus capture? *Developmental Science, 21*(6), e12667. https://doi.org/10.1111/desc.12667

Wass, S. V., Noreika, V., Georgieva, S., Clackson, K., Brightman, L., Nutbrown, R., … Leong, V. (2018). Parental neural responsivity to infants' visual attention: How mature brains influence immature brains during social interaction. *PLoS Biology*, *16*(12). https://doi.org/10.1371/journal.pbio.2006328

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

Webb, S. J., Long, J. D., & Nelson, C. A. (2005). A longitudinal investigation of visual event-related potentials in the first year of life. *Developmental Science, 8*(6), 605-616. https://doi.org/10.1111/j.1467-7687.2005.00452.x

Weber, A., & Varela, F. J. (2002). Life after Kant: Natural purposes and the autopoietic foundations of biological individuality. *Phenomenology and the Cognitive Sciences, 1*(2), 97–125. https://doi.org/10.1023/A:1020368120174

Weighall, A. R., Henderson, L. M., Barr, D. J., Cairney, S. A., & Gaskell, M. G. (2017). Eye-tracking the time-course of novel word learning and lexical competition in adults and children. *Brain and Language*, *167*, 13–27. https://doi.org/10.1016/j.bandl.2016.07.010

Weiss, S., & Rappelsberger, P. (2000). Long-range EEG synchronization during word encoding correlates with successful memory performance. *Cognitive Brain Research*, *9*(3), 299–312. https://doi.org/10.1016/S0926-6410(00)00011-2

Wellman, H. M., Cross, D., Bartsch, K., & Harris, P. L. (1986). Infant search and object permanence: A meta-analysis of the A-not-B error. *Monographs of the Society for Research in Child Development*, *51*(3). https://doi.org/10.2307/1166103

Wetherford, M. J., & Cohen, L. B. (1973). Developmental changes in infant visual preferences for novelty and familiarity. *Child Development*, *44*(3), 416–424. https://doi.org/10.2307/1127994

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin and Review, 9*(4), 625-636. https://doi.org/10.3758/BF03196322

Wolf, I., Dziobek, I., & Heekeren. (2010). Neural correlates of social cognition in naturalistic settings: A model-free analysis approach. *NeuroImage*, *49*(1), 894–904. https://doi.org/10.1016/j.neuroimage.2009.08.060

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*(1), 1–34. https://doi.org/10.1016/S0010-0277(98)00058-4

Woodward, A. L. (2003). Infants' developing understanding of the link between looker and object. *Developmental Science*, *6*(3), 297–311. https://doi.org/10.1111/1467-7687.00286

Woodward, A. L., Markman, E. M., & Fitzsimmons, C. M. (1994). Rapid word learning in 13- and 18-month-olds. *Developmental Psychology*, *30*(4), 553–566. https://doi.org/10.1037/0012-1649.30.4.553

Wynn, K. (1992). Addition and subtraction by human infants. *Nature, 358*(6389), 749-750. https://doi.org/10.1038/358749a0

Yip, J. A., & Schweitzer, M. E. (2015). Trust promotes unethical behavior: Excessive trust, opportunistic exploitation, and strategic exploitation. *Current Opinion in Psychology, 6*, 216-220. https://doi.org/10.1016/j.copsyc.2015.09.017

Yoon, J. M. D., Johnson, M. H., & Csibra, G. (2008). Communication-induced memory biases in preverbal infants. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(36), 13690–13695. https://doi.org/10.1073/pnas.0804388105

Yoshida, K. A., Pons, F., Maye, J., & Werker, J. F. (2010). Distributional phonetic learning at 10 months of age. *Infancy*, *15*(4), 420–433. https://doi.org/10.1111/J.1532-7078.2009.00024.X

Yu, C., & Smith, L. B. (2010). What you learn is what you see: Using eye movements to study infant cross-situational word learning. *Developmental Science*, *14*, 165–180. https://doi.org/ 10.1111/j.1467-7687.2010.00958.x

Yu, C., & Smith, L. B. (2012). Modeling cross-situational word-referent learning: Prior questions. *Psychological Review*, *119*(1), 21–39. https://doi.org/10.1037/A0026182

Zaheer, A., McEvily, B., & Perrone, V. (1998). Does trust matter? exploring the effects of interorganizational and interpersonal trust on performance. *Organization Science, 9*(2), 141-159. https://doi.org/10.1287/orsc.9.2.141

Zaki, J., & Ochsner, K. (2011). Reintegrating the study of accuracy into social cognition research. *Psychological Inquiry, 22*(3), 159-182. https://doi.org/10.1080/1047840X.2011.551743

Zmyj, N., Buttelmann, D., Carpenter, M., & Daum, M. M. (2010). The reliability of a model influences 14-month-olds' imitation. *Journal of Experimental Child Psychology*, *106*(4), 208–220. https://doi.org/10.1016/j.jecp.2010.03.002

Appendix: Chapter 2 supplementary materials

Table S1. The list of search terms used for the meta-analysis.

S1	**Search terms (PsycINFO, MEDLINE):** (hyperscan* OR dual-scan* OR "dual scan*" OR multi-scan* OR “multi scan*” OR multiple-scan* OR “multiple scan*”)
S2	**Search terms (PsycINFO, MEDLINE):** (((simultaneous* OR concurrent*) N8 (record* OR monitor* OR scan* OR collect* OR measure*)) AND (EEG* OR electroencephalogra* OR MEG* OR magnetoencephalogra* OR fNIRS* OR near-infrared spectroscop* OR fMRI*OR “functional magnetic resonance imag*” OR functional-magnetic-resonance-imag*))
S3	**Search terms (PsycINFO, MEDLINE):** ("dual EEG*" OR dual-EEG* OR "dual electroencephalogra*" OR dual-electroencephalogra* OR "dual MEG*" OR dual-MEG* OR "dual magnetoencephalogra*" OR dual-magnetoencephalogra* OR "dual fNIRS*" OR dual-fNIRS* OR “dual functional near-infrared spectroscop*" OR “dual-functional near-infrared spectroscop*" OR “dual fMRI*” OR dual-fMRI* OR “dual-functional magnetic resonance imag*” OR “dual functional magnetic resonance imag*” OR "multi EEG*" OR multi-EEG* OR "multiple EEG*" OR multiple-EEG* OR "multi electroencephalogra*" OR multi-electroencephalogra* OR “multiple electroencephalogra*" OR multiple-electroencephalogra* OR "multi MEG*" OR multi-MEG* OR "multiple MEG*" OR multiple-MEG* OR "multi magnetoencephalogra*" OR multi- magnetoencephalogra* OR “multiple magnetoencephalogra*" OR multiple- magnetoencephalogra* OR "multi fNIRS*" OR multi-fNIRS* OR “multi functional near-infrared spectroscop*" OR “multi-functional near-infrared spectroscop*" OR "multiple fNIRS*" OR multiple-fNIRS* OR “multiple functional near-infrared spectroscop*" OR “multiple-functional near-infrared spectroscop*" OR “multi fMRI*” OR multi-fMRI* OR “multi-functional magnetic resonance imag*” OR “multi functional magnetic resonance imag*” OR “multiple fMRI*” OR multiple-fMRI* OR “multiple-functional magnetic resonance imag*” OR “multiple functional magnetic resonance imag*”)
S4	**Search terms (PsycINFO, MEDLINE):** ((simultaneous* OR concurrent*) N8 (EEG* OR electroencephalogra* OR MEG* OR magnetoencephalogra* OR fNIRS* OR “near-infrared spectroscop*” OR fMRI* OR “functional magnetic resonance imag*” OR functional-magnetic-resonance-imag*))
S5	**Search terms (PsycINFO, MEDLINE):** ((“second person*” OR “second-person*” OR two-person* OR “two person*” OR “two brain*” or two-brain*) N3 neuroscien*)

S6	**Search terms (PsycINFO, MEDLINE):** ((inter-brain* OR interbrain* OR “inter brain*”OR between-brain* OR “between brain*” OR brain-to-brain* OR “brain to brain*” OR cross-brain* OR “cross brain*” OR two-brain* OR “two brain*” OR multi-brain* OR “multi brain*” OR multiple-brain* OR “multiple brain*”) N5 (coupled* OR coupling* OR coordinat* OR connect* OR synchron* OR align* OR entrain* OR shared* OR sharing* OR similar* OR network*))
S7	**Search terms (PsycINFO, MEDLINE):** ((inter-person* OR interperson* OR “inter person*” OR inter-subject* OR intersubject* OR “inter subject*” OR “two person*” OR two-person* OR “multi person*” OR multi-person* OR “multiple person*” OR multiple-person*) N5 (brain* OR neural*) N5 (coupled* OR coupling* OR coordinat* OR connect* OR synchron* OR align* OR entrain* OR shared* OR sharing* OR similar* OR network*))
S8	S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7

List S1. Review papers to be used for bibliographic search.

Babiloni, F., and Astolfi, L. (2014). Social neuroscience and hyperscanning techniques: past, present and future. *Neurosci. Biobehav. Rev. 44*, 76–93. https://doi.org/10.1016/j.neubiorev.2012.07.006

Bekkering, H., De Bruijn, E. R., Cuijpers, R. H., Newman-Norlund, R., Van Schie, H. T., & Meulenbroek, R. (2009). Joint action: Neurocognitive mechanisms supporting human interaction. *Topics in Cognitive Science*, *1*(2), 340-352. https://doi.org/10.1111/j.1756-8765.2009.01023.x

Czeszumski, A., Eustergerling, S., Lang, A., Menrath, D., Gerstenberger, M., Schuberth, S., ... & König, P. (2020). Hyperscanning: a valid method to study neural inter-brain underpinnings of social interaction. Frontiers in Human Neuroscience, *14*(39). 1-17. https://doi.org/10.3389/fnhum.2020.00039

Dumas, G., Lachat, F., Martinerie, J., Nadel, J., and George, N. (2011). From social behaviour to brain synchronization: review and perspectives in hyperscanning. IRBM, *32*(1), 48–53. https://doi.org/10.1016/j.irbm.2011.01.002

Hasson, U., Ghazanfar, A. A., Galantucci, B., Garrod, S., & Keysers, C. (2012). Brain-to-brain coupling: a mechanism for creating and sharing a social world. *Trends in cognitive sciences*, *16*(2), 114-121. https://doi.org/10.1016/j.tics.2011.12.007

Hoehl, S., Fairhurst, M., & Schirmer, A. (2020). Interactional synchrony: signals, mechanisms and benefits. Social Cognitive and Affective Neuroscience. *15*(1-2), 5-18. https://doi.org/10.1093/scan/nsaa024

Leong, V., & Schilbach, L. (2019). The promise of two-person neuroscience for developmental psychiatry: using interaction-based sociometrics to identify disorders of social interaction. *The British Journal of Psychiatry*, *215*(5), 636-638. https://doi.org/10.1192/bjp.2019.73

Markova, G., Nguyen, T., & Hoehl, S. (2019). Neurobehavioral interpersonal synchrony in early development: The role of interactional rhythms. *Frontiers in psychology*, *10*, 2078. https://doi.org/10.3389/fpsyg.2019.02078

Nguyen, T., Bánki, A., Markova, G., & Hoehl, S. (2020). Studying parent-child interaction with hyperscanning. *Progress in brain research, 254*, 1-24. https://doi.org/10.1016/bs.pbr.2020.05.003

Redcay, E., & Schilbach, L. (2019). Using second-person neuroscience to elucidate the mechanisms of social interaction. *Nature Reviews Neuroscience*, *20*(8), 495-505. https://doi.org/10.1038/s41583-019-0179-4

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., et al. (2013). Toward a second-person neuroscience. *Behav. Brain Sci. 36,* 393–414. https://doi.org/10.1017/s0140525x1200204x

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

Figure S1. PRISMA Flow diagram illustrating the process of systematic literature search and selection of eligible papers.

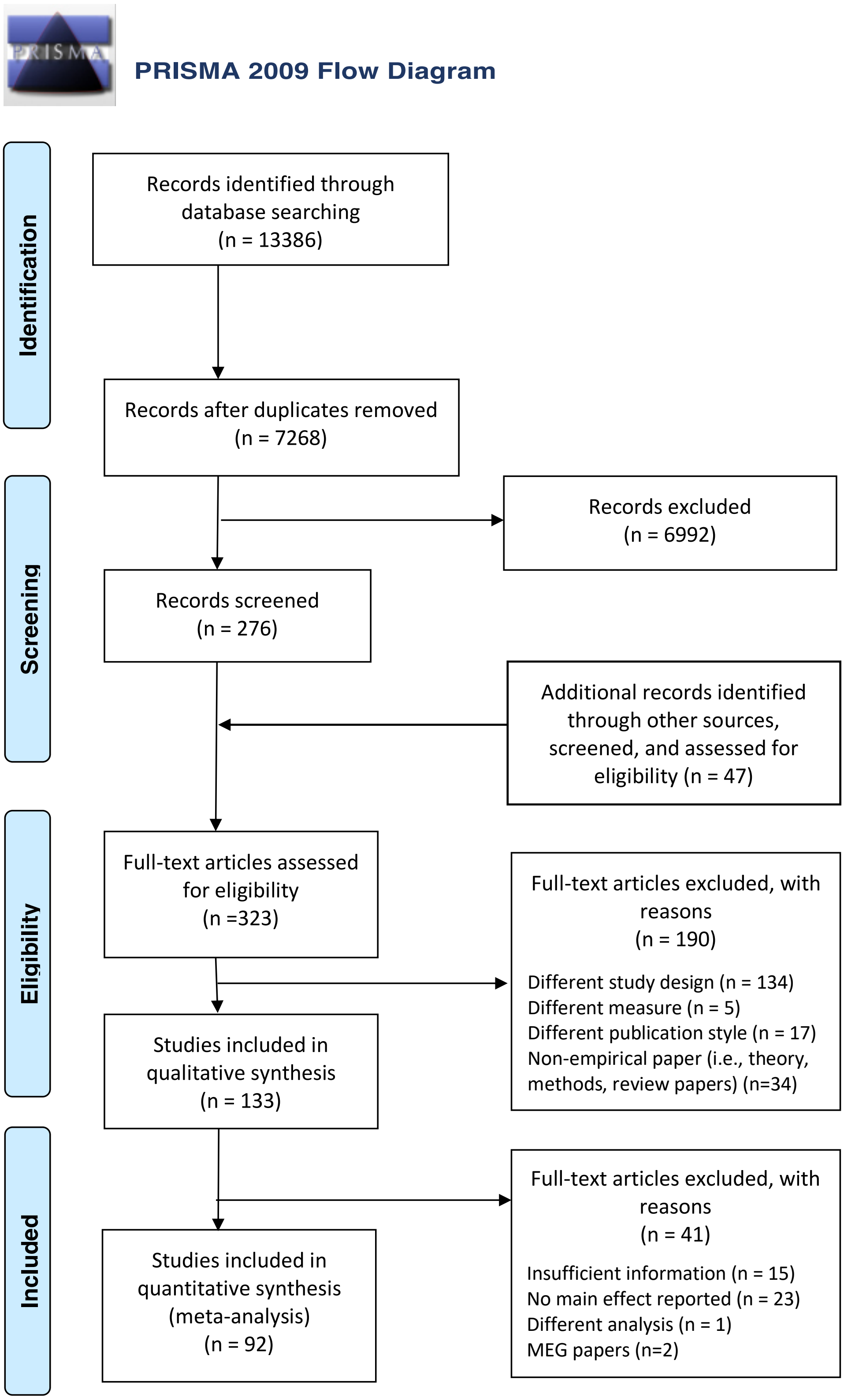


Table S2. Results of the full-text screening (papers excluded).

Reasons for full-text exclusion	Description and/or examples
Different study design (*n*=134)	• No hyperscanning (i.e., only one brain scanned at a time) • Data collected on a non-interactional setting (e.g., watching the same visual stimuli with no interaction with other participants) • Research question focusing on identifying neural characteristics of certain group rather than clarifying neural interaction between two or more people - Typically, these studies use the term inter-personal/subject correlation/coherence as an analytic method to measure a common response to external stimuli, rather than a measure reflecting an association of two or more brain activities. Hence, such studies often perform analysis according to, for instance, certain characteristics of participants (e.g., role in the task, the presence/absence of clinical diagnosis) and involve no analysis on dyad level. • Used brain stimulation (e.g., tACS) - Studies utilising brain stimulation were excluded because such data are not comparable to data of non-stimulated brain activities collected during a naturalistic situation. • Used neurofeedback - A study involving neurofeedback was excluded because the situation was considered not as naturalistic to the participants as other studies included, and hence the data was considered not comparable to the data reported in other studies.
Different measure (*n*=5)	• Non-brain data collected and/or analysed (e.g., heart rate)
Different publication style (*n*=17)	• Not an original research article with empirical data (e.g., “Journal club”)
Methods paper (*n*=22)	• Papers whose main aim was to propose a novel paradigm or analytical pipeline. • Papers with no empirical data were reported for the method validation (e.g., simulated data were used) • Papers with data collected in one condition and no statistical analysis reported (i.e., no effect size reported).
Theory paper (*n*=9)	• Papers whose aim was to demonstrate a theoretical framework of hyperscanning research or brain-to-brain interaction. • Commentary to theoretical papers.
Review paper (*n*=3)	• Review papers on the topic of hyperscanning technique or brain-to-brain interaction.

List S2. List of EEG papers included in the meta-analysis (Chapter 2).

[1] Antonenko, P. D., Davis, R., Wang, J., & Celepkolu, M. (2019). On the same wavelength: Exploring team neurosynchrony in undergraduate dyads solving a cyberlearning problem with collaborative scripts. *Mind, Brain, and Education*, *13*(1), 4-13. https://doi.org/10.1111/mbe.12187

[2] Astolfi, L., Toppi, J., De Vico Fallani, F., Vecchiato, G., Salinari, S., Mattia, D., … Babiloni, F. (2010). Neuroelectrical hyperscanning measures simultaneous brain activity in humans. *Brain Topography*, *23*(3), 243–256. https://doi.org/10.1007/s10548-010-0147-9

[3] Balconi, M., Gatti, L., & Vanutelli, M. E. (2018). EEG functional connectivity and brain-to-brain coupling in failing cognitive strategies. *Consciousness and Cognition: An International Journal*, *60*, 86–97. https://doi.org/10.1016/j.concog.2018.03.001

[4] Balconi, M., & Vanutelli, M. E. (2018). Functional EEG connectivity during competition. *BMC Neuroscience*, *19*(1). https://doi.org/10.1186/s12868-018- 0464-6

[5] Bridwell, D. A., Henderson, S., Sorge, M., Plis, S., & Calhoun, V. D. (2018). Relationships between alpha oscillations during speech preparation and the listener N400 ERP to the produced speech. *Scientific Reports*, *8*(1), 12838. https://doi.org/10.1038/s41598-018-31038-9

[6] Ciaramidaro, A., Toppi, J., Casper, C., Freitag, C. M., Siniatchkin, M., & Astolfi, L. (2018). Multiple-brain connectivity during third party punishment: an EEG hyperscanning study. *Scientific Reports*, *8*(1), 6822. https://doi.org/10.1038/s41598-018-24416-w

[7] Dumas, G., Nadel, J., Soussignan, R., Martinerie, J., & Garnero, L. (2010). Inter-brain synchronization during social interaction. *PloS One*, *5*(8), e12166. https://doi.org/10.1371/journal.pone.0012166

[8] Hinterberger, T., Von Haugwitz, A., & Schmidt, S. (2016). Does a healing procedure referring to theta rhythms also generate theta rhythms in the Brain? *Journal of Alternative and Complementary Medicine*, *22*(1), 66–74. https://doi.org/10.1089/acm.2014.0394

[9] Hu, Y., Pan, Y., Shi, X., Cai, Q., Li, X., & Cheng, X. (2018). Inter-brain synchrony and cooperation context in interactive decision making. *Biological Psychology*, *133*, 54–62. https://doi.org/10.1016/j.biopsycho.2017.12.005

[10] Jahng, J., Kralik, J. D., Hwang, D. U., & Jeong, J. (2017). Neural dynamics of two players when using nonverbal cues to gauge intentions to cooperate during the Prisoner's Dilemma Game. *NeuroImage*, *157*, 263–274. https://doi.org/10.1016/j.neuroimage.2017.06.024

[11] Kinreich, S., Djalovski, A., Kraus, L., Louzoun, Y., & Feldman, R. (2017). Brain-to-brain synchrony during naturalistic social interactions. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-17339-5

[12] Leong, V., Byrne, E., Clackson, K., Georgieva, S., Lam, S., & Wass, S. (2017). Speaker gaze increases information coupling between infant and adult brains. *Proceedings of the National Academy of Sciences, 114*(50), 13290-13295. https://doi.org/10.1073/pnas.1702493114

[13] Liu, H., Li, B., Wang, X., & He, Y. (2020). Role of joint language control during cross-language communication: evidence from cross-frequency coupling. *Cognitive Neurodynamics, 15*, 191-205. https://doi.org/10.1007/s11571-020-09594-6

[14] Mu, Y., Guo, C., & Han, S. (2016). Oxytocin enhances inter-brain synchrony during social coordination in male adults. *Social Cognitive and Affective Neuroscience*, *11*(12), 1882–1893. https://doi.org/10.1093/scan/nsw106

[15] Mu, Y., Han, S., & Gelfand, M. J. (2017). The role of gamma interbrain synchrony in social coordination when humans face territorial threats. *Social Cognitive and Affective Neuroscience*, *12*(10), 1614–1623. https://doi.org/10.1093/scan/nsx093

[16] Müller, V., & Lindenberger, U. (2014). Hyper-brain networks support romantic kissing in humans. *PloS One*, *9*(11), e112080. https://doi.org/10.1371/journal.pone.0112080

[17] Pérez, A., Dumas, G., Karadag, M., & Duñabeitia, J. A. (2019). Differential brain-to-brain entrainment while speaking and listening in native and foreign languages. *Cortex*, *111*, 303–315. https://doi.org/10.1016/j.cortex.2018.11.026

[18] Santamaria, L., Noreika, V., Georgieva, S., Clackson, K., Wass, S., & Leong, V. (2020). Emotional valence modulates the topology of the parent-infant inter-brain network. *NeuroImage*, *207*. 116341. https://doi.org/10.1016/j.neuroimage.2019.116341

[19] Sun, H., Verbeke, W. J. M. I. M. I., Pozharliev, R., Bagozzi, R. P., Babiloni, F., & Wang, L. (2019). Framing a trust game as a power game greatly affects interbrain synchronicity between trustor and trustee. *Social Neuroscience*, *14*(6), 635–648. https://doi.org/10.1080/17470919.2019.1566171

[20] van Vugt, M. K., Pollock, J., Johnson, B., Gyatso, K., Norbu, N., Lodroe, T., … Fresco, D. M. (2020). Inter-brain synchronization in the practice of tibetan monastic debate. *Mindfulness*, *11*(5), 1105–1119. https://doi.org/10.1007/s12671-020-01338-1

[21] Wang, C., Li, H., Jia, L., Li, F., & Wang, J. (2020). Theta band behavioral fluctuations synchronized interpersonally during cooperation. *Psychonomic Bulletin & Review*, *27*(3), 563–570. https://doi.org/10.3758/s13423-020-01711-0

[22] Yun, K., Watanabe, K., & Shimojo, S. (2012). Interpersonal body and neural synchronization as a marker of implicit social interaction. *Scientific Reports*, *2*(959), 959. https://doi.org/10.1038/srep00959

[23] Zhang, D., Lin, Y., Jing, Y., Feng, C., & Gu, R. (2019). The dynamics of belief updating in human cooperation: Findings from inter-brain ERP hyperscanning. *NeuroImage*, *198*, 1–12. https://doi.org/10.1016/j.neuroimage.2019.05.029

List S3. List of fNIRS papers included in the meta-analysis (Chapter 2).

[1] Baker, J. M., Liu, N., Cui, X., Vrticka, P., Saggar, M., Hosseini, S. M. H. H., & Reiss, A. L. (2016). Corrigendum: Sex differences in neural and behavioral signatures of cooperation revealed by fNIRS hyperscanning. *Scientific Reports*, *6*, 30512. https://doi.org/10.1038/srep30512

[2] Balconi, M., & Fronda, G. (2020). The “gift effect” on functional brain connectivity. Inter-brain synchronization when prosocial behavior is in action. *Scientific Reports*, *10*(1), 5394. https://doi.org/10.1038/s41598-020-62421-0

[3] Balconi, M., Gatti, L., & Vanutelli, M. E. (2018). When cooperation goes wrong: Brain and behavioural correlates of ineffective joint strategies in dyads. *International Journal of Neuroscience*, *128*(2), 155–166. https://doi.org/10.1080/00207454.2017.1379519

[4] Balconi, M., Pezard, L., Nandrino, J.-L. L., & Vanutelli, M. E. (2017). Two is better than one: The effects of strategic cooperation on intra- and inter-brain connectivity by fNIRS. *PloS One*, *12*(11), e0187652. https://doi.org/10.1371/journal.pone.0187652

[5] Balconi, M., & Vanutelli, M. E. (2017a). Brains in competition: Improved cognitive performance and inter-brain coupling by hyperscanning paradigm with functional near-infrared spectroscopy. *Frontiers in Behavioral Neuroscience*, *11*. https://doi.org/10.3389/fnbeh.2017.00163

[6] Balconi, M., & Vanutelli, M. E. (2017b). Interbrains cooperation: Hyperscanning and self-perception in joint actions. *Journal of Clinical and Experimental Neuropsychology*, *39*(6), 607–620. https://doi.org/10.1080/13803395.2016.1253666

[7] Balconi, M., Vanutelli, M. E., & Gatti, L. (2018). Functional brain connectivity when cooperation fails. *Brain and Cognition*, *123*, 65–73. https://doi.org/10.1016/j.bandc.2018.02.009

[8] Cheng, X., Li, X., & Hu, Y. (2015). Synchronous brain activity during cooperative exchange depends on gender of partner: A fNIRS-based hyperscanning study. *Human Brain Mapping*, *36*(6), 2039–2048. https://doi.org/10.1002/hbm.22754

[9] Cheng, X., Pan, Y., Hu, Y., & Hu, Y. (2019). Coordination elicits synchronous brain activity between co-actors: Frequency ratio matters. *Frontiers in Neuroscience*, *13*, 1071. https://doi.org/10.3389/fnins.2019.01071

[10] Cui, X., Bryant, D. M., & Reiss, A. L. (2012). NIRS-based hyperscanning reveals increased interpersonal coherence in superior frontal cortex during cooperation. *NeuroImage*, *59*(3), 2430–2437. https://doi.org/10.1016/j.neuroimage.2011.09.003

[11] Dai, B., Chen, C., Long, Y., Zheng, L., Zhao, H., Bai, X., … Lu, C. (2018). Neural mechanisms for selectively tuning in to the target speaker in a naturalistic noisy situation. *Nature Communications*, *9*(1), 2405. https://doi.org/10.1038/s41467-018-04819-z

[12] Dai, R., Liu, R., Liu, T., Zhang, Z., Xiao, X., Sun, P., … Zhu, C. (2018). Holistic cognitive and neural processes: A fNIRS-hyperscanning study on interpersonal sensorimotor synchronization. *Social Cognitive and Affective Neuroscience*, *13*(11), 1141–1154. https://doi.org/10.1093/scan/nsy090

[13] Feng, X., Sun, B., Chen, C., Li, W., Wang, Y., Zhang, W., … Shao, Y. (2020). Self–other overlap and interpersonal neural synchronization serially mediate the effect of behavioral synchronization on prosociality. *Social Cognitive and Affective Neuroscience*, *15*(2), 203–214. https://doi.org/10.1093/scan/nsaa017

[14] Fishburn, F. A., Murty, V. P., Hlutkowsky, C. O., MacGillivray, C. E., Bemis, L. M., Murphy, M. E., … Perlman, S. B. (2018). Putting our heads together: Interpersonal neural synchronization as a biological mechanism for shared intentionality. *Social Cognitive and Affective Neuroscience*, *13*(8), 841–849. https://doi.org/10.1093/scan/nsy060

[15] Hirsch, J., Zhang, X., Noah, J. A., & Ono, Y. (2017). Frontal temporal and parietal systems synchronize within and across brains during live eye-to-eye contact. *NeuroImage*, *157*, 314–330. https://doi.org/10.1016/j.neuroimage.2017.06.018

[16] Holper, L., Scholkmann, F., & Wolf, M. (2012). Between-brain connectivity during imitation measured by fNIRS. *NeuroImage*, *63*(1), 212–222. https://doi.org/10.1016/j.neuroimage.2012.06.028

[17] Hou, Y., Song, B., Hu, Y., Pan, Y., & Hu, Y. (2020). The averaged inter-brain coherence between the audience and a violinist predicts the popularity of violin performance. *NeuroImage*, *211*, 116655. https://doi.org/10.1016/j.neuroimage.2020.116655

[18] Hu, Y., Hu, Y., Li, X., Pan, Y., & Cheng, X. (2017). Brain-to-brain synchronization across two persons predicts mutual prosociality. *Social Cognitive and Affective Neuroscience*, *12*(12), 1835–1844. https://doi.org/10.1093/scan/nsx118

[19] Ikeda, S., Nozawa, T., Yokoyama, R., Miyazaki, A., Sasaki, Y., Sakaki, K., & Kawashima, R. (2017). Steady beat sound facilitates both coordinated group walking and inter-subject neural synchrony. *Frontiers in Human Neuroscience*, *11*, 147. https://doi.org/10.3389/fnhum.2017.00147

[20] Jiang, J., Chen, C., Dai, B., Shi, G., Ding, G., Liu, L., … Fiske, S. T. (2015). Leader emergence through interpersonal neural synchronization. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, *112*(14), 4274–4279. https://doi.org/10.1073/pnas.1422930112

[21] Jiang, J., Dai, B., Peng, D., Liu, L., Lu, C., Zhu, C., … Lu, C. (2012). Neural synchronization during face-to-face communication. *The Journal of Neuroscience*, *32*(45), 16064–16069. https://doi.org/10.1523/JNEUROSCI.2926-12.2012

[22] Koide, T., & Shimada, S. (2018). Cheering enhances inter-brain synchronization between sensorimotor areas of player and observer. *Japanese Psychological Research*, *60*(4), 265–275. https://doi.org/10.1111/jpr.12202

[23] Kruppa, J. A., Reindl, V., Gerloff, C., Weiss, E. O., Prinz, J., Herpertz-Dahlmann, B., … Schulte-Rüther, M. (2020). Brain and motor synchrony in children and adolescents with ASD- an fNIRS hyperscanning study. *Social Cognitive and Affective Neuroscience, 16*(1-2), 103-116. https://doi.org/10.1093/scan/nsaa092

[24] Li, L., Wang, H., Luo, H., Zhang, X., Zhang, R., & Li, X. (2020). Interpersonal neural synchronization during Cooperative behavior of basketball players: A fNIRS-based hyperscanning study. *Frontiers in Human Neuroscience*, *14*, 169. https://doi.org/10.3389/fnhum.2020.00169

[25] Li, Y., Chen, R., Turel, O., Feng, T., Zhu, C.-Z. Z., & He, Q. (2020). Dyad sex composition effect on inter-brain synchronization in face-to-face cooperation. *Brain Imaging and Behavior, 15*(3), 1667-1675. https://doi.org/10.1007/s11682-020-00361-z

[26] Liu, J., Zhang, R., Geng, B., Zhang, T., Yuan, D., Otani, S., & Li, X. (2019). Interplay between prior knowledge and communication mode on teaching effectiveness: Interpersonal neural synchronization as a neural marker. *NeuroImage*, *193*, 93–102. https://doi.org/10.1016/j.neuroimage.2019.03.004

[27] Liu, N., Mok, C., Witt, E. E., Pradhan, A. H., Chen, J. E., & Reiss, A. L. (2016). NIRS-based hyperscanning reveals inter-brain neural synchronization during cooperative Jenga game with face-to-face communication. *Frontiers in Human Neuroscience*, *10*, 82. https://doi.org/10.3389/fnhum.2016.00082

[28] Liu, T., Saito, H., & Oi, M. (2015). Role of the right inferior frontal gyrus in turn-based cooperation and competition: A near-infrared spectroscopy study. *Brain and Cognition*, *99*, 17–23. https://doi.org/10.1016/j.bandc.2015.07.001

[29] Liu, W., Branigan, H. P., Zheng, L., Long, Y., Bai, X., Li, K., … Lu, C. (2019). Shared neural representations of syntax during online dyadic communication. *NeuroImage*, *198*, 63–72. https://doi.org/10.1016/j.neuroimage.2019.05.035

[30] Lu, K., & Hao, N. (2019). When do we fall in neural synchrony with others? *Social Cognitive and Affective Neuroscience*, *14*(3), 253–261. https://doi.org/10.1093/scan/nsz012

[31] Lu, K., Qiao, X., & Hao, N. (2019). Praising or keeping silent on partner's ideas: Leading brainstorming in particular ways. *Neuropsychologia*, *124*, 19–30. https://doi.org/10.1016/j.neuropsychologia.2019.01.004

[32] Lu, K., Xue, H., Nozawa, T., & Hao, N. (2019). Cooperation makes a group be more creative. *Cerebral Cortex*, *29*(8), 3457–3470. https://doi.org/10.1093/cercor/bhy215

[33] Lu, K., Yu, T., & Hao, N. (2020). Creating while taking turns, the choice to unlocking group creative potential. *NeuroImage*, *219*, 117025. https://doi.org/10.1016/j.neuroimage.2020.117025

[34] Mayseless, N., Hawthorne, G., & Reiss, A. L. (2019). Real-life creative problem solving in teams: fNIRS based hyperscanning study. *NeuroImage*, *203*, 116161. https://doi.org/10.1016/j.neuroimage.2019.116161

[35] Miller, J. G., Vrtička, P., Cui, X., Shrestha, S., Hosseini, S. M. H., Baker, J. M., & Reiss, A. L. (2019). Inter-brain synchrony in mother-child dyads during cooperation: An fNIRS hyperscanning study. *Neuropsychologia*, *124*, 117–124. https://doi.org/10.1016/j.neuropsychologia.2018.12.021

[36] Nguyen, T., Schleihauf, H., Kayhan, E., Matthes, D., Vrtička, P., & Hoehl, S. (2020a). Neural synchrony in mother-child conversation: Exploring the role of conversation patterns. *Social Cognitive and Affective Neuroscience, 16*(1-2), 93-102. https://doi.org/10.1093/scan/nsaa079

[37] Nguyen, T., Schleihauf, H., Kayhan, E., Matthes, D., Vrtička, P., & Hoehl, S. (2020b). The effects of interaction quality on neural synchrony during mother-child problem solving. *Cortex*, *124*, 235–249. https://doi.org/10.1016/j.cortex.2019.11.020

[38] Nozawa, T., Sakaki, K., Ikeda, S., Jeong, H., Yamazaki, S., Kawata, K. H. dos S., … Kawashima, R. (2019). Prior physical synchrony enhances rapport and inter-brain synchronization during subsequent educational communication. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-49257-z

[39] Pan, Y., Cheng, X., Zhang, Z., Li, X., & Hu, Y. (2017). Cooperation in lovers: An fNIRS- based hyperscanning study. *Human Brain Mapping*, *38*(2), 831–841. https://doi.org/10.1002/hbm.23421

[40] Pan, Y., Dikker, S., Goldstein, P., Zhu, Y., Yang, C., & Hu, Y. (2020). Instructor-learner brain coupling discriminates between instructional approaches and predicts learning. *NeuroImage*, *211*, 116657. https://doi.org/10.1016/j.neuroimage.2020.116657

[41] Pan, Y., Guyon, C., Borragán, G., Hu, Y., & Peigneux, P. (2020). Interpersonal brain synchronization with instructor compensates for learner’s sleep deprivation in interactive learning. *Biochemical Pharmacology*, 114111. https://doi.org/10.1016/j.bcp.2020.114111

[42] Pan, Y., Novembre, G., Song, B., Li, X., & Hu, Y. (2018). Interpersonal synchronization of inferior frontal cortices tracks social interactive learning of a song. *NeuroImage*, *183*, 280–290. https://doi.org/10.1016/j.neuroimage.2018.08.005

[43] Piazza, E. A., Hasenfratz, L., Hasson, U., & Lew-Williams, C. (2020). Infant and adult brains are coupled to the dynamics of natural communication. *Psychological Science*, *31*(1), 6–17. https://doi.org/10.1177/0956797619878698

[44] Piva, M., Zhang, X., Noah, J. A., Chang, S. W. C. C., & Hirsch, J. (2017). Distributed neural activity patterns during human-to-human competition. *Frontiers in Human Neuroscience*, *11*, 571. https://doi.org/10.3389/fnhum.2017.00571

[45] Quiñones-Camacho, L. E., Fishburn, F. A., Camacho, M. C., Hlutkowsky, C. O., Huppert, T. J., Wakschlag, L. S., & Perlman, S. B. (2020). Parent–child neural synchrony: a novel approach to elucidating dyadic correlates of preschool irritability. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *61*(11), 1213–1223. https://doi.org/10.1111/jcpp.13165

[46] Reindl, V., Gerloff, C., Scharke, W., & Konrad, K. (2018). Brain-to-brain synchrony in parent-child dyads and the relationship with emotion regulation revealed by fNIRS-based hyperscanning. *NeuroImage*, *178*, 493–502. https://doi.org/10.1016/j.neuroimage.2018.05.060

[47] Sun, B., Xiao, W., Feng, X., Shao, Y., Zhang, W., & Li, W. (2020). Behavioral and brain synchronization differences between expert and novice teachers when collaborating with students. *Brain and Cognition*, *139*, 105513. https://doi.org/10.1016/j.bandc.2019.105513

[48] Tang, H., Mai, X., Wang, S., Zhu, C., Krueger, F., & Liu, C. (2016). Interpersonal brain synchronization in the right temporo-parietal junction during face-to-face economic exchange. *Social Cognitive and Affective Neuroscience*, *11*(1), 23–32. https://doi.org/10.1093/scan/nsv092

[49] Tang, Y., Liu, X., Wang, C., Cao, M., Deng, S., Du, X., … Li, F. (2020). Different strategies, distinguished cooperation efficiency, and brain synchronization for couples: An fNIRS-based hyperscanning study. *Brain and Behavior, 10*(9), e01768. https://doi.org/10.1002/brb3.1768

[50] Wang, Q., Han, Z., Hu, X., Feng, S., Wang, H., Liu, T., & Yi, L. (2020). Autism symptoms modulate interpersonal neural synchronization in children with autism spectrum disorder in cooperative interactions. *Brain Topography*, *33*(1), 112–122. https://doi.org/10.1007/s10548-019-00731-x

[51] Xu, J., Slagle, J. M., Banerjee, A., Bracken, B., & Weinger, M. B. (2019). Use of a portable functional near-infrared spectroscopy (fNIRS) system to examine team experience during crisis event management in clinical simulations. *Frontiers in Human Neuroscience*, *13*, 85. https://doi.org/10.3389/fnhum.2019.00085

[52] Xue, H., Lu, K., & Hao, N. (2018). Cooperation makes two less-creative individuals turn into a highly-creative pair. *NeuroImage*, *172*, 527–537. https://doi.org/10.1016/j.neuroimage.2018.02.007

[53] Zhang, M., Jia, H., Zheng, M., & Liu, T. (2021). Group decision-making behavior in social dilemmas: Inter-brain synchrony and the predictive role of personality traits. *Personality and Individual Differences*, *168*, 110315. https://doi.org/10.1016/j.paid.2020.110315

[54] Zhang, M., Liu, T., Pelowski, M., Jia, H., & Yu, D. (2017). Social risky decision-making reveals gender differences in the TPJ: A hyperscanning study using functional near-infrared spectroscopy. *Brain and Cognition*, *119*, 54–63. https://doi.org/10.1016/j.bandc.2017.08.008

[55] Zhang, M., Liu, T., Pelowski, M., & Yu, D. (2017). Gender difference in spontaneous deception: A hyperscanning study using functional near-infrared spectroscopy. *Scientific Reports*, *7*(1), 1-13. https://doi.org/10.1038/s41598-017-06764-1

[56] Zhang, Y., Meng, T., Hou, Y., Pan, Y., & Hu, Y. (2018). Interpersonal brain synchronization associated with working alliance during psychological counseling. *Psychiatry Research: Neuroimaging*, *282*, 103–109. https://doi.org/10.1016/j.pscychresns.2018.09.007

[57] Zheng, L., Chen, C., Liu, W., Long, Y., Zhao, H., Bai, X., … Lu, C. (2018). Enhancement of teaching outcome through neural prediction of the students' knowledge state. *Human Brain Mapping*, *39*(7), 3046–3057. https://doi.org/10.1002/hbm.24059

[58] Zheng, L., Liu, W., Long, Y., Zhai, Y., Zhao, H., Bai, X., … Lu, C. (2020). Affiliative bonding between teachers and students through interpersonal synchronisation in brain activity. *Social Cognitive and Affective Neuroscience*, *15*(1), 97–109. https://doi.org/10.1093/scan/nsaa016

List S4. List of fMRI papers included in the meta-analysis (Chapter 2).

[1] Abe, M. O., Koike, T., Okazaki, S., Sugawara, S. K., Takahashi, K., Watanabe, K., & Sadato, N. (2019). Neural correlates of online cooperation during joint force production. *NeuroImage*, *191*, 150–161. https://doi.org/10.1016/j.neuroimage.2019.02.003

[2] Bilek, E., Ruf, M., Schäfer, A., Akdeniz, C., Calhoun, V. D., Schmahl, C., … Meyer-Lindenberg, A. (2015). Information flow between interacting human brains: Identification, validation, and relationship to social expertise. *PNAS*, *112*(16), 5207–5212. https://doi.org/10.1073/pnas.1421831112

[3] Dikker, S., Silbert, L. J., Hasson, U., & Zevin, J. D. (2014). On the same wavelength: Predictable language enhances speaker–listener brain-to-brain synchrony in posterior superior temporal gyrus. *The Journal of Neuroscience*, *34*(18), 6267–6272. https://doi.org/10.1523/JNEUROSCI.3796-13.2014

[4] Kerr, K. L., Cosgrove, K. T., Ratliff, E. L., Burrows, K., Misaki, M., Moore, A. J., … Morris, A. S. (2020). TEAMwork: Testing emotional attunement and mutuality during parent-adolescent fMRI. *Frontiers in Human Neuroscience*, *14*, 24. https://doi.org/10.3389/fnhum.2020.00024

[5] King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R., & Montague, P. R. (2005). Getting to know you: Reputation and trust in a two-person economic exchange. *Science, 308*(5718), 78-83. https://doi.org/10.1029/2003PA000992

[6] Koike, T., Tanabe, H. C., Adachi-Abe, S., Okazaki, S., Nakagawa, E., Sasaki, A. T., … Sadato, N. (2019). Role of the right anterior insular cortex in joint attention-related identification with a partner. *Social Cognitive and Affective Neuroscience*, *14*(10), 1131–1145. https://doi.org/10.1093/scan/nsz087

[7] Koike, T., Tanabe, H. C., Okazaki, S., Nakagawa, E., Sasaki, A. T., Shimada, K., … Sadato, N. (2016). Neural substrates of shared attention as social memory: A hyperscanning functional magnetic resonance imaging study. *NeuroImage*, *125*, 401–412. https://doi.org/10.1016/j.neuroimage.2015.09.076

[8] Kostorz, K., Flanagin, V. L., & Glasauer, S. (2020). Synchronization between instructor and observer when learning a complex bimanual skill. *NeuroImage*, *216*, 116659. https://doi.org/10.1016/j.neuroimage.2020.116659

[9] Shaw, D. J., Czekóová, K., Špiláková, B., Salazar, M., Řezáč, P., Kurečková, V., … Brázdil, M. (2020). A neuroscientific evaluation of driver rehabilitation: Functional neuroimaging demonstrates the effectiveness of empathy induction in altering brain responses during social information processing. *PLoS ONE*, *15*(4), e0232222. https://doi.org/10.1371/journal.pone.0232222

[10] Shaw, D. J., Czekóová, K., Staněk, R., Mareček, R., Urbánek, T., Špalek, J., … Brázdil, M. (2018). A dual-fMRI investigation of the iterated Ultimatum Game reveals that reciprocal behaviour is associated with neural alignment. *Scientific Reports*, *8*(1), 1-13. https://doi.org/10.1038/s41598-018-29233-9

[11] Špiláková, B., Shaw, D. J., Czekóová, K., Mareček, R., & Brázdil, M. (2020). Getting into sync: Data-driven analyses reveal patterns of neural coupling that distinguish among different social exchanges. *Human Brain Mapping*, *41*(4), 1072–1083. https://doi.org/10.1002/hbm.24861

[12] Stolk, A., Noordzij, M. L., Verhagen, L., Volman, I., Schoffelen, J. M., Oostenveld, R., … Toni, I. (2014). Cerebral coherence between communicators marks the emergence of meaning. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(51), 18183–18188. https://doi.org/10.1073/pnas.1414886111

Table S3. Results of meta-regression analyses taking no intercept to calculate estimates of overall effect sizes for each subgroup. Note. K refers to the number of analyses included in the subgroup. Estimate denotes the overall effect size of the subgroup, with the SE being the standard error of the estimate, and the p-value, and 95% confidence interval associated to the estimated overall effect size.

Sample size

		k	estimate	*SE*	*p*	low 95% *CI*	high 95% *CI*
EEG	Moderators	*F*(1,130) = 4.24			.041		
	Residual Heterogeneity	*QE*(*df* = 130) = 266.18			< .0001		
	Intercept	132	1.26	0.24	< .0001	0.80	1.73
	Sample size		-0.02	0.01	.041	-0.04	0.00
fNIRS	Moderators	*F*(1,227) = 5.69			.018		
	Residual Heterogeneity	*QE*(*df* = 227) = 312.89			< .0001		
	Intercept	229	1.34	0.14	< .0001	1.07	1.62
	Sample size		-0.01	0.00	.018	-0.01	0.00
fMRI	Moderators	*F*(1,78) = 0.20			.66		
	Residual Heterogeneity	*QE*(*df* = 78) = 111.10			.0082		
	Intercept	80	0.69	0.63	.28	-0.57	1.95
	Sample size		0.01	0.03	.66	-0.04	0.07

Participant Age

		k	estimate	*SE*	*p*	low 95% *CI*	high 95% *CI*
EEG	Moderators	F(2,130) = 21.5477			< .0001		
	Residual Heterogeneity	*QE*(*df* = 130) = 305.10			< .0001		
	Adult	99	0.86	0.14	< .0001	0.59	1.13
	Adult-Child	33	0.84	0.41	.044	0.02	1.66
fNIRS	Moderators	*F*(2,227) = 75.6317			< .0001		
	Residual Heterogeneity	*QE*(*df* = 227) = 321.93			< .0001		
	Adult	182	1.13	0.10	< .0001	0.93	1.31
	Adult-Child	48	0.85	0.23	.0002	0.41	1.30
fMRI	Moderators	*F*(2,78) = 8.70			.0004		
	Residual Heterogeneity	*QE*(*df* = 78) = 111.10			.0017		
	Adult	79	0.88	0.24	.0005	0.40	1.36
	Adult-Child	1	2.58	1.26	.045	0.06	5.09

Gender							
						low 95% *CI*	high 95% *CI*
		k	estimate	*SE*	*p*		
EEG	Moderators	*F*(4,128) =12.1830			< .0001		
	Residual Heterogeneity	*QE*(*df* = 128) = 253.37			< .0001		
	Same	64	0.83	0.16	<.0001	0.51	1.15
	Opposite	9	0.55	0.37	.1347	-0.17	1.28
	Mixed	48	0.80	0.30	.0083	0.21	1.39
	Not reported	11	1.48	0.40	.0003	0.69	2.28
fNIRS	Moderators	*F*(5,224) = 35.819			< .0001		
	Residual Heterogeneity	*QE*(*df* = 224) = 284.07			.0040		
	Same	73	1.10	0.13	<.0001	0.85	1.35
	Opposite	15	2.08	0.34	<.0001	1.41	2.74
	Mixed	105	0.86	0.14	<.0001	0.59	1.13
	Manipulated	6	0.94	0.38	.014	0.19	1.69
	Not reported	31	1.22	0.22	<.0001	0.79	1.65
fMRI	Moderators	*F*(3,77) = 114.55			.0011		
	Residual Heterogeneity	*QE*(*df* = 77) = 114.55			.0036		
	Same	76	0.79	0.26	.0031	0.27	1.31
	Mixed	3	1.91	0.73	.011	0.45	3.37
	Not reported	1	1.46	1.15	.208	-0.83	3.75

Relationship							
		k	estimate	*SE*	*p*	low 95% *CI*	high 95% *CI*
EEG	Moderators	*F*(3,128) = 0.3522			.79		
	Residual Heterogeneity	*QE*(*df* = 128) = 292.33			< .0001		
	Close	34	0.85	0.38	.028	0.09	1.60
	Strangers	68	0.92	0.18	< .0001	0.55	1.28
	Manipulated	2	0.36	0.61	.67	-0.94	1.46
	Not reported	28	0.87	0.24	.0004	0.40	1.28
fNIRS	Moderators	*F*(5,225) = 29.8962			< .0001		
	Residual Heterogeneity	*QE*(*df* = 224) = 318.57			< .0001		
	Close	44	1.07	0.18	< .0001	0.71	1.43
	Strangers	129	0.96	0.12	< .0001	0.73	1.20
	Mixed	3	1.22	0.66	.0650	0.36	1.36
	Manipulated	8	0.86	0.25	.0008	0.36	1.36
	Not reported	46	0.19	0.19	< .0001	1.08	1.82
fMRI	Moderators	*F*(3,77) = 5.39			.002		
	Residual Heterogeneity	*QE*(*df* = 77) = 91.25			.13		
	Close	1	2.58	1.29	.049	0.01	5.14
	Strangers	45	0.84	0.32	.012	0.19	1.48
	Not reported	34	0.96	0.41	.022	0.14	1.77

Interaction Medium							
		k	estimate	*SE*	*p*	low 95% *CI*	high 95% *CI*
EEG	Moderators	*F*(4.128) = 11.52			< .0001		
	Residual Heterogeneity	*QE*(*df* = 128) = 241.6			< .0001		
	Computer	39	0.92	0.15	< .0001	0.62	1.21
	Direct	76	0.91	0.14	< .0001	0.63	1.19
	Limited	1	0.95	0.76	.22	-0.56	2.45
	Manipulated	16	0.39	0.39	.32	-0.38	1.17
fNIRS	Moderators	*F*(5,224) = 38.395			< .0001		
	Residual Heterogeneity	*QE*(*df* = 224) = 275.11			.0110		
	Computer	35	1.15	0.17	< .0001	0.81	1.49
	Direct	168	1.00	0.10	< .0001	0.80	1.22
	Limited	8	3.14	0.60	< .0001	1.97	4.31
	Manipulated	18	0.92	0.21	< .0001	0.50	1.40

Interaction Modality							
		k	estimate	*SE*	*p*	low 95% *CI*	high 95% *CI*
EEG	Moderators	*F*(5,127) = 8.65			< .0001		
	Residual Heterogeneity	*QE*(*df* = 127) = 208.81			< .0001		
	Visual only	10	0.81	0.35	.02	0.12	1.51
	Auditory only	15	0.75	0.47	.11	-0.18	1.68
	Visual and auditory	62	0.93	0.22	< .0001	0.49	1.37
	Computer	29	0.99	0.24	< .0001	0.51	1.47
	Manipulated	16	0.40	0.41	.34	-0.42	1.21
fNIRS	Moderators	*F*(5,224) = 28.35			< .0001		
	Residual Heterogeneity	*QE*(*df* = 224) = 316.63			< .0001		
	Visual only	36	1.22	0.25	< .0001	0.71	1.71
	Auditory only	1	0.65	0.72	.37	-0.77	2.06
	Visual and auditory	140	1.08	0.13	< .0001	0.83	1.34
	Computer	34	1.16	0.20	< .0001	0.78	1.55
	Manipulated	19	0.94	0.24	.0001	0.47	1.41
fMRI	Moderators	*F*(4,76) = 4.58			.0023		
	Residual Heterogeneity	*QE*(df = 76) = 79.17			.38		
	Visual only	49	1.12	0.29	.0002	0.54	1.70
	Auditory only	3	1.24	0.68	.073	-0.12	2.60
	Motor	1	0.25	0.89	.78	-1.52	2.02
	Visual and motor	27	0.04	0.68	.95	-1.32	1.40

Cognitive Function							
		k	estimate	*SE*	*p*	low 95% *CI*	high 95% *CI*
EEG	Moderators	*F*(6,126) = 6.48			< .0001		
	Residual Heterogeneity	*QE*(*df* = 126) = 222.78			< .0001		
	Action coordination	22	0.68	0.31	.0290	0.07	1.29
	Competition vs Cooperation	6	1.40	0.65	.0334	0.11	2.69
	Game	23	0.85	0.29	.0036	0.28	1.41
	Membership	1	1.40	0.87	.1121	-0.33	3.13
	Speech and Communication	66	0.85	0.24	.0005	0.38	1.32
	Representation	14	0.89	0.32	.0056	0.27	1.52
fNIRS	Moderators	*F*(6,223) = 25.8470			< .0001		
	Residual Heterogeneity	*QE*(*df* = 223) = 294.14			.0010		
	Action coordination	98	1.07	0.15	< .0001	0.78	1.36
	Competition vs Cooperation	13	1.87	0.39	< .0001	1.10	2.64
	Game	29	0.98	0.20	< .0001	0.59	1.37
	Representation	6	1.08	0.46	.0188	0.18	1.98
	Speech and Communication	83	1.06	0.14	< .0001	0.77	1.34
	Manipulated	1	0.61	0.69	.3744	-0.74	1.96
fMRI	Moderators	*F*(5,75) = 5.0042			.0005		
	Residual Heterogeneity	*QE*(*df* = 75) = 50.28			.9900		
	Coordination	27	0.04	0.58	.9434	-1.11	1.20
	Competition vs Cooperation	1	0.07	0.73	.9190	-1.37	1.52
	Game	24	0.86	0.34	.0142	0.18	1.54
	Speech and Communication	3	1.25	0.63	.0516	-0.01	2.51
	Representation	25	1.55	0.40	.0002	0.75	2.36

Measure of Inter-brain Coupling							
		k	estimate	*SE*	*p*	low 95% *CI*	high 95% *CI*
EEG	Moderators	*F*(4,128) = 9,7860			< .0001		
	Residual Heterogeneity	*QE*(df = 128) = 223.55			< .0001		
	Coupling	57	0.83	0.19	< .0001	0.46	1.20
	Correlation	27	0.89	0.28	.0018	0.34	1.44
	Information flow	29	1.11	0.39	.0049	0.34	1.88
	Graph index	19	0.76	0.38	.0491	0.00	1.52
fNIRS	Moderators	*F*(4,225) = 35			< .0001		
	Residual Heterogeneity	*QE*(*df* = 225) = 323.86			< .0001		
	Coupling	189	1.05	0.10	< .0001	0.85	1.25
	Correlation	27	1.26	0.22	< .0001	0.82	1.70
	Information flow	13	1.11	0.22	< .0001	0.69	1.54
	Psychophysiological interaction analysis	1	0.88	0.87	.3100	-0.82	2.59
fMRI	Moderators	*F*(2,78) = 7.91			.0007		
	Residual Heterogeneity	*QE*(*df* = 78) = 194.26			< .0001		
	Correlation	75	0.82	0.26	.0020	0.31	1.33
	Information Flow	5	1.30	0.55	.0202	020	2.37

Statistical Comparison							
		k	estimate	*SE*	*p*	low 95% *CI*	high 95% *CI*
EEG	Moderators	*F*(4,128) = 12.43			< .0001		
	Residual Heterogeneity	*QE*(*df* = 128) = 207.05			< .0001		
	AB	89	0.75	0.15	< .0001	0.46	1.04
	AC	27	0.96	0.27	.0005	0.43	1.49
	Correlational	1	0.95	0.74	.2056	-0.53	2.42
	Surrogate	15	1.27	0.35	.0003	0.59	1.96
fNIRS	Moderators	*F*(4,225) = 35.60			< .0001		
	Residual Heterogeneity	*QE*(*df* = 225) = 325.92			< .0001		
	AB	107	1.09	0.11	< .0001	0.88	1.30
	AC	55	0.93	0.15	< .0001	0.64	1.22
	Chance level	58	1.29	0.14	< .0001	1.02	1.56
	Surrogate	10	1.24	0.25	< .0001	0.75	1.73
fMRI	Moderators	*F*(4,76) = 6.61			.0001		
	Residual Heterogeneity	*QE*(*df* = 76) = 30.50			.0000		
	AB	11	0.99	0.35	.0065	0.29	1.70
	AC	12	1.24	0.45	.0076	0.34	2.14
	Chance level	19	1.66	0.38	< .0001	0.91	2.42
	Surrogate	38	0.55	0.28	.0507	0.00	1.10

Figure S5. Forest plots (EEG studies).

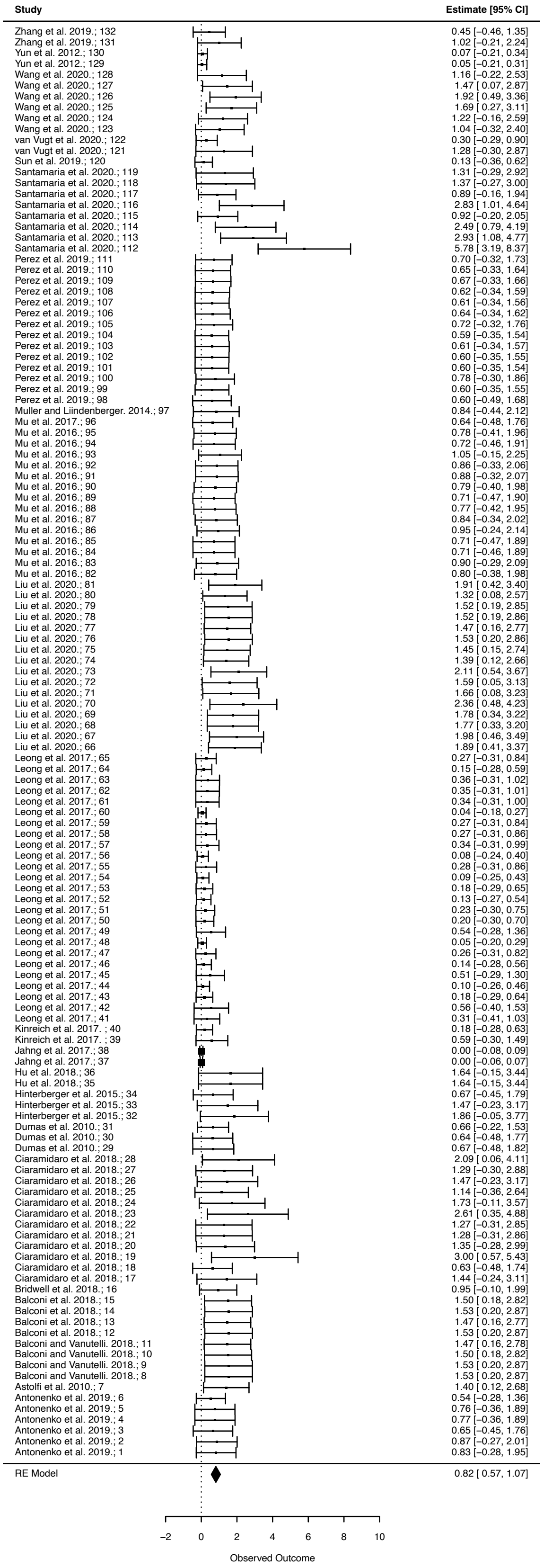


Figure S6. Forest plots (fNIRS studies).

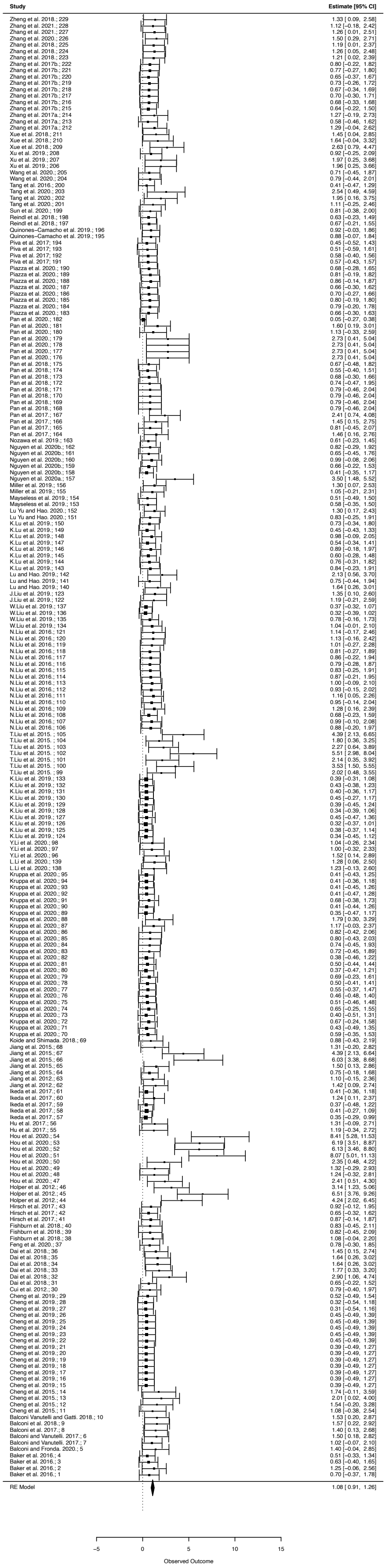


Figure S7. Forest plots (fMRI studies).