THE MULTIFACETED NATURE OF HUMAN INTERACTION

by

ROKSANA MARKIEWICZ

A thesis submitted to
The University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

School of Psychology
College of Life and Environmental Sciences
University of Birmingham
September 2023

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive
e-theses repository

This unpublished thesis/dissertation is under a Creative Commons Attribution 4.0 International (CC BY 4.0) licence.

You are free to:

Share — copy and redistribute the material in any medium or format

Adapt — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Notices:

You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation.

No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material.

Unless otherwise stated, any material in this thesis/dissertation that is cited to a third-party source is not included in the terms of this licence. Please refer to the original source(s) for licencing conditions of any quotes, images or other material cited to a third party.

ABSTRACT

In everyday life, we engage in interactions with others to express our thoughts and desires, employing verbal and non-verbal communication methods. This thesis examined the multifaceted nature of effective communication and joint action. **Chapter 1** provides an overview of relevant literature related to language, cooperation, and joint action as well as relevant background information about the electroencephalography (EEG) method used in my thesis. In **Chapter 2**, using EEG and a minimal two-word paradigm, we investigated age-related oscillatory mechanisms of lexical retrieval and semantic binding. Results reported a different and delayed age-related oscillatory signature for semantic binding, suggesting that older adults rely on different mechanisms and take longer to integrate the word meaning into the semantic context. **Chapter 3** examined the brain-to-brain neural coupling that underlie the mechanisms of developing shared representations of a specific task goal in response to feedback. We utilised EEG hyperscanning and a newly developed non-verbal cooperation paradigm. We found that an anti-correlation of theta power between two individuals following negative feedback forecasted successful convergence of the representation of the task goal and thus successful cooperation. Conversely, correlated activity in the theta band forecasted a lack of convergence and subsequent cooperative failure. Lastly, **Chapter 4** considered the impact of theory of mind (i.e. ToM) abilities on verbal cooperation in a novel verbal two-player paradigm. We found that pairs featuring two individuals with high ToM abilities committed less cooperative errors and subsequently cooperated better compared to pairs where both individuals have low ToM abilities. Overall, the findings from this thesis highlight the complex nature of human interaction and offer new perspectives on its study.

ACKNOWLEDGMENTS

None of this would have been possible without the support of some of the most wonderful and encouraging people. For this, I will always be grateful <here is me attempting to make a hand symbol for half of a heart – but the Gen Z way>.

Firstly, a special thank you to Katrien Segaert and Ali Mazaheri - the best supervisory duo to exist. Ali – you believed in me from the start. During my masters you said to me: “but you’ll do a PhD, right?” After this, I went home and googled what a PhD was… I will forever be grateful for everything you taught me, EEGwise and beyond, your endless support, encouragement in moments when I needed it the most, and for injecting your love for science in me. Katrien – thank you for teaching me not to be scared to problem solve, it turns out I am pretty good at it now ;) Thank you for always wiping away my tears, offering endless hugs, but also for acknowledging when life was tough. I will forever be grateful for your life (and PhD) advice, for teaching me how to do good science and overcoming obstacles. Having you both on my team has truly shaped this PhD journey.

I would also like to thank Andrea, with whom I started my research journey. Thank you for helping me realise my ambition to become a researcher, for that I will be forever grateful.

I would like to show my gratitude to the Hilary Green studentship for funding this project.

Emma – I am so glad we did this PhD journey together. Thank you for all the laughs we shared, all the coffees we drank, for rubber ducking, for bearing with my aggressive (!) driving, for making me drink water in the office, for making us dinners when we had no time to eat, for helping with painting my walls. Thank you for being the best PhD partner. Look at us – all grown up now!

To Marion, who has always been present throughout my time at UoB. Whether it is for science advice, a night of good food and board games, or excellent cake, you were always there, thank you! To Christopher – for all your witty jokes. It still takes me a while to work out if you’re serious or joking!

To Sammy – my fellow PhD on the other side of the world. The phone catch-ups always made me feel like there was someone who understood me. Thank you for always telling me/ us: ‘We will be ok’. And we made it!

To the PMC – Sophie, Zheni and Ceci. Thanks for being the older, (sometimes) wiser committee. I always look forward to our chats when you try to convince me to jump ship! I am still holding on strong… for now.

To my office mates (and adopted office mates!): Rupali and Alice – our Friday gossip clubs kept me going through each week. Thank you for all the stories, entertainment and all the laughs. Alice – I hope all the cooking/ baking was worth this acknowledgment! To Amélie, Monica, and Salma for always being a listening ear. To Yulia, fellow pickle lover, for never making me feel alone when turning up to a quiz night ‘just to sit there and look pretty’.

To Melka, you were one of the most important parts of my life, I am forever grateful for all the memories. You were so special. You will always have a piece of my heart.

And finally, but most importantly, thank you to the boyos – S & G. Sami - thank you for always being so kind to me. Thank you for making the best dinners, listening to me, debugging my programming problems, brainstorming with me, being interested in my work (or at least pretending!), making me laugh, and loving me (even at my weirdest points). Thank you for celebrating all my little successes, but mostly thank you for picking up the pieces when things didn’t go so well. This work really wouldn’t have been possible without you. You are my safe place. I can’t wait for our next adventures together!

LIST OF PUBLICATIONS

Below lists published work (in peer-reviewed journals, pre-print servers, and at conferences) included in this thesis and outlines author contributions for each:

Chapter 2: The entire body of the work presented in this chapter was formed from the below publication.

Markiewicz, R., Segaert, K., & Mazaheri, A. (2021). How the healthy ageing brain supports semantic binding during language comprehension. *European Journal of Neuroscience, 54*(11), 7899–7917. https://doi.org/10.1111/ejn.15525

Author contributions: KS and AM conceptualized the study. RM collected the data and analysed the data. AM and KS supervised data analyses. RM and KS wrote the manuscript. All authors edited the manuscript.

Chapter 3: The entire body of the work presented in this chapter was formed from the below publication.

Markiewicz, R., Segaert, K., & Mazaheri, A. Brain to brain coupling forecasts future joint action outcomes (Manuscript submited for peer-review; pre-print available: https://doi.org/10.1101/2023.09.21.558775).

Author contributions: RM, KS and AM conceptualised the study. RM collected the data and analysed the data. AM and KS supervised data analyses. RM wrote the manuscript. All authors edited the manuscript.

Chapter 4: The entire body of the work presented in this chapter was formed from the below publication.

Markiewicz, R., Rahman, F., Apperly, I., Mazaheri, A., Segaert, K., (2023). It Is Not All About You: Communicative Cooperation Is Determined by Your Partner's Theory of Mind Abilities as Well as Your Own. Journal of Experimental Psychology : Learning , Memory , and Cognition. https://doi.org/https://doi.org/10.1037/xlm0001268

Author contributions: RM, KS, and AM conceptualized the study with advice from IA. RM programmed the experiment. RM collected and analyzed the data. KS supervised data analyses. RM and FR wrote the manuscript. All authors contributed to data interpretation, drafting, revising, and final approval of the manuscript.

Media coverage

Markiewicz, R. (2023, July 13). Want to collaborate better? Pick your partner wisely and learn how to read them. *The Conversation*. Retrieved from https://theconversation.com/.

Conference Presentations

Markiewicz, R., Segaert, K., & Mazaheri, A. (2023). Society for the Neurobiology of Language 15th Annual Meeting, Marseille, France

Markiewicz, R., Segaert, K., & Mazaheri, A. (2023). British Association for Cognitive Neuroscience, Cardiff, UK

Markiewicz, R., Segaert, K., & Mazaheri, A. (2023). Research Day, Birmingham, UK.

Markiewicz, R., Segaert, K., & Mazaheri, A. (2021). The BRAIN conference (online).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: GENERAL INTRODUCTION AND THESIS OUTLINE

Effective interactions with others are fundamental in the society. On a daily basis, we interact with others in order to articulate our thoughts, ideas and needs. This transpires through the utilization of ***language***, encompassing both spoken and written forms. Successful communication is not solely confined to verbal means; it can also manifest through non-verbal channels such as gestures, gaze, or adaptive behavioural adjustments in response to input received, for example in the context of ***joint actions***. The success of communication or cooperation is reliant on social cognition, specifically the mentalizing processes such as ***theory of mind***, which facilitate the understanding of each other (Frith & Frith, 2012).

The overarching goal of this thesis is to explore the multifaceted aspects of effective cooperation, communication, and joint action. The unifying thread weaving through all experimental chapters is the pivotal intersection of language and human interaction. One way of studying human interaction is via adopting a micro-level perspective and focusing on the specific linguistic processes (for example the construction of minimal phrases), which serve as a foundation for human interaction. With this groundwork, we can then zoom out and study human interaction beyond the linguistic properties of language. By doing so, we gain a more comprehensive understanding of the broader concepts of social communication.

Specifically, the two original aims of the thesis were as follows:

- To examine the brain-to-brain neural patterns that underlie the mechanisms of developing shared representations of a task goal (i.e., successful cooperation) in response to feedback in a non-verbal cooperation task (**Chapter 3**).
- To examine individual differences, specifically Theory of Mind abilities, that may influence cooperative success in a verbal cooperation task (**Chapter 4**).

Originally, in addition to the above, my thesis aimed to investigate the influence of Theory of Mind capabilities on the success of cooperation in *non-verbal* collaborative contexts. However, the emergence of the COVID-19 pandemic necessitated alterations to my research plans as in-person human data collection was restricted throughout 2020 and 2021. The nature of the non-verbal cooperation task requiring precise timing measurements (as exact reaction times were of interest) precluded its adaptation to an online platform. In addition, due to time constraints and the necessary large number of participants due to the *individual differences* type of the study, it was not feasible to pursue this original third objective of the thesis. Instead, I transitioned the study focusing on *verbal* cooperation and Theory of Mind (**Chapter 4**) into an online format allowing for the acquisition of valuable data within the limitations imposed by the pandemic. Secondly, in order to optimise productivity during the pandemic I analysed an existing dataset (**Chapter 2**) with the following aim:

- To investigate the differences in modulations of oscillatory brain activity related to language processing, specifically lexical retrieval and semantic binding, between young and healthy older adults (**Chapter 2**).

By doing so, I acquired all the necessary EEG analytical skills, enabling me to transition to collecting and analysing complex EEG hyperscanning data (**Chapter 3**) upon the resumption of in-person research activities in late 2021.

This introductory chapter establishes a comprehensive groundwork for the subsequent sections, navigating through key methodologies and concepts employed in the thesis. Firstly, I introduce the EEG technique and its traditional use in investigating cognitive processes, providing a foundational understanding essential for interpreting the findings in Chapter 2, which delves into language comprehension. Subsequently, I elaborate on its novel application in the study of social interactions, aiming for heightened ecological validity, as detailed in

Chapter 3. I then offer a comprehensive review of relevant literature for each experimental chapter, while elucidating the complexities of language processing and human interaction, which is a theme resonating throughout all experimental chapters. The core exploration of social communication, interaction and language intricacies permeates the thesis, with specific aspects studied in individual experimental chapters. Thus I provide background literature and present key concepts encompassing (a) semantic binding and lexical retrieval in both young and older adults, (b) the dynamics of joint action, prediction of others' actions and adaptive responses to feedback, and (c) ToM and its link to cooperative behaviour. This introductory chapter serves as a foundation guide for the thesis, and the exploration of the multifaceted nature of human interaction.

1.1 Introduction to electroencephalography (EEG)

Considering the overarching utilization of the electroencephalography (EEG) methodology within this thesis, coupled with its innovative implementation through EEG hyperscanning, one primary objective of the introductory section is to provide a comprehensive background and overview of the EEG technique. This will encompass its fundamental operational principles, and its potential for advancing EEG research towards a heightened ecological validity, particularly in the realm of developing more authentic and real-life social interactions within the EEG laboratory framework.

1.1.1 EEG as a conventional tool

Nerve cells communicate with one another and the rest of the body by sending electrical signals. EEG records the summed electrical activity stemming from excitatory and inhibitory post-synaptic potentials over ensembles (thousands/ millions) of neurons on the surface of the scalp. This results in rhythmic patterns of activity which form neuronal oscillations.

While EEG offers millisecond-level accuracy, its temporal precision surpasses its spatial resolution, making precise localization of brain regions challenging, especially in instances involving signals originating from deep within the brain (Cohen, 2014b; Curham & Allen, 2022). Research that utilises the EEG technique often examines how the EEG signal changes in response to an event (often stimuli appearing on the screen). This can be divided into evoked and induced activity. The evoked signal is both time- and phase-locked to an event, with the assumption that a large proportion of the EEG activity is unrelated to the event. Instead, the event-related signal (i.e., event-related potential; ERP) is extracted from the EEG signal, centred around the experimental event and averaged across several trials. The peaks and troughs of the ERP waveform form components, with each of them being related to a task-relevant cognitive process (Mazaheri, 2022). In language processing research, one of the most prominent and famous components is the N400 component, a negative deflection peaking at around between 250-400ms post experimental event, linked to processing of semantics (Kutas & Hillyard, 1980).

On the other hand, the induced signal is time-locked but not necessarily phase locked to the event at hand and relies on the postulation that the experimental event modulates the on-going brain activity. This approach involves averaging the time-frequency spectra over several trials, also centred at the experimental event (Mazaheri, 2022). Neuronal oscillations (or time-frequency spectra) contain information encompassing frequency, power (or amplitude), and phase. Frequency pertains to the speed of oscillation, quantified by the cycles completed per second. Power signifies the energy magnitude within a specific frequency band. Phase delineates the oscillation's position on a sine wave at a particular time point (Cohen, 2014b). Oscillatory activity can be split into approximate frequency bands: theta (4-7Hz), alpha (8-14Hz), beta (15-30Hz), and gamma (30-100Hz), which map onto various brain regions and cognitive functions. Activity in the theta band is thought to be related to the processing of cognitive demands; in language processing studies this reflects the processing of semantic or syntactic violations (Davidson & Indefrey, 2007; Hald et al., 2006; Wang et al., 2012). Alpha band activity over the occipital areas has famously been linked to attentional processing (Jensen & Mazaheri, 2010), whereas in the scope of language processing, low beta has been linked to semantic and syntactic binding (Bastiaansen et al., 2010; Bastiaansen & Hagoort, 2006; von Stein et al., 1999; Weiss & Mueller, 2003).

1.1.2 EEG in context of hyperscanning

The EEG technique has recently been used in a novel way, aiming to bridge the gap between neuroimaging laboratory research and real-life contextual settings. Social interaction has long posed challenges for ecologically valid measurement using neuroimaging methods. This is where EEG hyperscanning, along with other neuroimaging modalities, emerges as a potential transformative method. EEG hyperscanning involves the simultaneous measurement of (neuroelectric) brain activity from two or more individuals. This technique facilitates examination into how the dynamic activity patterns of two or more brains contribute to the ongoing adjustment of one's actions in reaction to the evolving behaviours of another individual. This technological advance has directed a paradigm shift in the study of social cognition, with numerous studies highlighting the emergence of 'synchrony' between brains in cooperative or joint action scenarios. This synchrony between individuals can manifest itself across motor, perceptual, or cognitive domains (Knoblich et al., 2011), with behavioural alignment being mirrored in the neural alignment. This neural alignment is denoted by the correlation between the temporally aligned neural signals, most often phase of an oscillation, but occasionally also power/ amplitude.

One theory that has recently emerged and holds promise in explaining the analogous patterns of brain activity between two interacting individuals is the Mutual prediction theory (Kingsbury et al., 2019). This theory postulates that within the context of social interactions, an individual’s brain not only encodes predictive information about their own behaviour but also anticipates and encodes predictive information about the behaviour of the other person. Thus the sum of activity of person’s A brain will be similar to the sum of activity in person’s B brain (Kingsbury et al., 2019). This provides insight into the observed synchrony during cooperation while its absence during competition (Cui et al., 2012). In cooperative interactions, both partners must engage in predicting each other's actions to act in tandem or in similar fashions, which may explain the neural synchrony. Conversely, competitive scenarios discourage mutual prediction of each other's actions, as speed and individual performance take precedence (Hamilton, 2021).

Hyperscanning has been used in prior research to explore a range of domains, encompassing social interaction (Balconi & Vanutelli, 2017b; Cui et al., 2012; Stolk et al., 2013), the dynamics of cooperation versus competition (Mu et al., 2016), musical engagement (Babiloni et al., 2012) , and even classroom dynamics (Dikker et al., 2017) (see Czeszumski et al., 2020; Konvalinka & Roepstorff, 2012 for reviews). Utilization of this technique is bringing neuroscience toward more ecologically valid experiments, where the intricate interplay of brain activity (and behaviour) involving multiple individuals is considered.

The conventional EEG methodology is used in **Chapter 2**, in which brain oscillatory activity is investigated in relation to language processing. The use of EEG hyperscanning is employed in **Chapter 3**, where a two-player non-verbal paradigm is developed to explore the brain-to-brain dynamics that forecast the cooperative outcomes following external feedback.

1.2 Human interaction and language

Human interaction, communication, and language are the core concepts examined throughout this thesis. Speaking is the most intricate cognitive-motor ability that humans possess (Indefrey & Levelt, 2000). However, effective communication and social interaction (i.e., exchange of meaning) in humans does not solely depend on language production (it is not all about the speaker!). Understanding the speaker's message requires the examination of the input and output structures. There are two main approaches for verbal communication; the unilateral and bilateral accounts (Clark & Krych, 2004). The unilateral account postulates that speaking and listening are separate processes; speakers autonomously shape their utterances, while listeners attempt to interpret these utterances. On the other hand, the bilateral account suggests that speaking and listening are joint processes, whereby the speakers take into account their own and their addresses actions during speech (Clark & Krych, 2004). Verbal communication as a bilateral process heavily relies on *grounding* – that is speakers strive to create mutual belief that their partner's understanding of them is sufficient enough for the current goals. The process of successful grounding in verbal communication relies on the following mechanisms: (1) *attention* to the speech, (2) *identification* of the presented utterances, (3) *understanding* of the meaning of the utterances by establishing common ground, and (4) *considering* answering the question of the speaker. In the bilateral account, the monitoring of the other occurs during each of these mechanisms. It has been suggested that other-monitoring grants most efficient grounding (Clark & Brennan, 1991). Clark and Krych (2004) provided evidence that when other-monitoring is not possible, there is a cost associated with compensation for the absence of such monitoring. Similarly, mechanisms like joint attention, the establishment of common ground, and the observation of actions are fundamental components within the concept of joint action (refer to Section 1.3 for more detail on joint action). These parallel mechanisms not only underscore the interconnected and multifaceted nature of human interaction, the central theme of this thesis, but also emphasize the shared foundation grounded in mutual understanding and alignment leading to successful human interaction. The components of language, communication, and social interaction partly intersect, together forming a rich network of shared cognitive processes and communicative dynamics. Each experimental chapter examines a different facet of human interaction and communication. **Chapter 2** examines the neural underpinnings of language processing, specifically, semantic binding in young and older adults. Chapter 3 explores the brain-to-brain dynamics underlying adjustment in a joint action scenario. Lastly, **Chapter 4** investigates the influence of individual differences on verbal bilateral communication.

1.2.1 Semantic binding

One unique aspect of human interaction and particularly of language is that we are able to merge words in novel ways in order to craft meaningful utterances. The significance of an individual word (e.g., flat) can be completely transformed by adding a subsequent word (e.g., flat tire vs. flat note). We possess the ability to build complex meaning from more elementary building blocks (Hagoort, 2020; Hagoort et al., 2009) – this is referred to as *semantic binding, unification* (Hagoort, 2005), or *merge* (Chomsky, 1995). In exploring these linguistic phenomena, we are able to zoom in on the specific language processes involved in human interaction, unravelling the facets that contribute to the richness of communication and social dynamics.

There are three functional aspects thought to form the essence of language processing (Hagoort, 2005). Firstly, *memory* – this pertains to the language related information stored within long term memory (i.e., mental lexicon) and the manner in which this information is accessed (i.e., lexical access). Secondly, *unification* or *binding* is a process where the lexically retrieved information is assembled into a more complex multiword utterance. Lastly, the *control* element establishes the connection between language and action. This comes into play when managing the exchange of turns in a conversation (Hagoort et al., 2009). Language processing but particularly semantic binding has been widely studied using neuroimaging techniques. It is essential to delve into the micro levels of linguistic processing (such as semantic binding) to draw a comprehensive picture of human interaction and communication dynamics.

1.2.2 Neural mechanisms of semantic binding in young adults

Given the incredibly fast rate of processing language (i.e., normal speech rate ranges from 120-200 words per minute (Liberman et al., 1967; Rayner, 1998), with an even faster reading rate of 250-350 words per minute (Rayner & Clifton, 2009)), considering temporal precision is crucial in order to understand the neural underpinnings of language processing and in turn human interaction. Therefore, the area of most interest here is research that utilised EEG due to its precise temporal resolution which allows to investigate the specifically timed neural mechanisms that underlie language processing. Previous studies investigating how the brain processes semantic comprehension have primarily used the ERP approach. It has consistently been found that young adults are able to integrate the word meaning into the meaning of a sentence level utterance at around 400ms following the target word (Kutas & Federmeier, 2011). This has been attributed to the N400 ERP component. The N400 was first reported by Kutas and Hillyard in 1980, maximal at the parietal sites and largest in response to semantic anomalies (but also present for improbable utterances), and later linked to semantic unification (Hagoort et al., 2009). Subsequent research began to study the N400 in more detail and found that the N400 is indeed related to semantic violations but is modulated by world knowledge (Hagoort et al., 2004). On the other hand, given that the N400 effects occur with pseudowords also, the N400 component might also reflect earlier language processing including orthographic and phonological stages (Deacon et al., 2004).

In addition to evoked activity, the EEG signal encompasses rhythmic oscillations. Brain oscillations play a crucial function in shaping how the brain perceives the surrounding environment, including the processing of language. Previous literature has shown the involvement of various frequency bands (of particular importance are theta and beta bands) in language processing (see reviews in Meyer, 2018; Prystauka & Lewis, 2019), with changes in different frequency bands mapping onto different language processes. For example, increased theta band power reflects an allocation of increased neural resources. In regard to language processing this theta power increase is often a result of integration of semantic anomalies (e.g., Bastiaansen & Hagoort, 2015; Davidson & Indefrey, 2007; Hald et al., 2006) and syntactic violations (e.g., Bastiaansen et al., 2002; Kielar et al., 2015; Lewis et al., 2016) into the broader context. Beta power oscillations also play a pivotal role in processing linguistic information. It has been shown that the beta power increase is a result of the required active maintenance of linguistic cues, specifically needed for complex sentence structures (Lewis et al., 2015; Lewis & Bastiaansen, 2015; Meyer et al., 2013). Conversely, beta power decreases are thought to reflect the need for processing change of the incoming linguistic information. Specifically, the presence of semantic and syntactic anomalies leads to beta power decreases compared with when these violations are not present (Kielar et al., 2014, 2015, 2018; Lewis & Bastiaansen, 2015; Luo et al., 2010; Meyer et al., 2013).

1.2.3 Neural mechanisms of semantic binding in older adults

Understanding the neural mechanisms underlying language processing in the young adult population holds crucial significance, as it lays the foundation for subsequent comparisons with older participants to discern the impact of aging on this cognitive function. Healthy ageing is accompanied by some declines in the cognitive abilities whilst others remain intact (behaviourally). Specifically, research has shown steady declines in cognitive ‘fluid’ abilities as we age including how fast we process information, working memory and memory for events (Salthouse, 1996, 2010; Waters & Caplan, 2005). On the other hand, cognitive abilities such as semantic memory and ‘fact knowing’ remain unchanged with age (Salthouse, 2010). The picture of language and healthy ageing remains complex. Our ability to retrieve word-related information for production declines as we age (Hardy et al., 2020;

Katrien Segaert, Lucas, et al., 2018). On the other hand older adults possess a greater vocabulary size (Brysbaert et al., 2016) and their sentence comprehension remains stable (Peelle, 2019; Shafto & Tyler, 2014) (unless the conditions become more difficult such as introducing background noise; Tun, 1998). Although (some) behavioural aspects of language processing thus remain unaltered with advancing age, it is plausible that the underlying processing mechanisms are orchestrated by distinct neural foundations.

Neuroimaging studies have reported nuanced disparities in the temporal dynamics and processing strategies of language between young and older healthy adults. The N400 effects are often diminished or/ and delayed in message-level congruity contexts in older adults compared to their young counterparts (Federmeier & Kutas, 2005; Wlotko & Federmeier, 2012). Additionally, older adults experience limitations in employing prediction mechanisms during language comprehension (Federmeier et al., 2002; Wlotko et al., 2012). Moreover, there is frequently observed diminished neural activity within regions linked with task related involvement among older adults, co-occurring with heightened engagement of alternative brain regions (Cabeza et al., 1997, 2002; Grady, 2000). Some studies have also found inverse oscillatory patterns in the alpha band specifically related to language processing compared to young adults (Beese et al., 2019; Poulisse et al., 2020). These findings collectively highlight the complex interplay of age, cognitive processes, and neural dynamics within the multifaceted domain of language comprehension.

While there exists, partial evidence illustrating distinctive neural patterns in language comprehension between older and younger adults, the predominant emphasis in previous research lies in examining ERPs within this context of aging. Thus, aiming to address this gap in the existing literature, **Chapter 2** aims to examine the modulations of oscillatory activity between young and healthy older individuals, in relation to processing of language comprehension (with a specific focus on semantic binding).

1.3 Joint action

Research into joint action and cooperation has experienced a recent surge in attention, this includes verbal and non-verbal cooperation, communication and human interaction. This expansion stems from the necessity to comprehensively encompass all participants engaged in these processes, inherently involving two or more individuals. Devising studies capable of capturing this dynamic while maintaining measurable independent variables and outcomes has proven challenging, owing to the intricate interplay between methodology and technology, particularly within the realm of neuroscience. As explained above, in the past decade, substantial technological advancements have propelled the emergence of two-brain (or, hyperscanning) research, with a significant expansion in the investigation of joint action. Throughout this thesis, the terms ‘joint action’ and ‘cooperation’ will be used interchangeably.

Joint action, whether it involves carrying a heavy sofa up a flight of stairs, dancing duets or more simply converging on a representation of an idea (or more arduously - depending on who you are converging with!), is present in everyday life. We as humans often manage to coordinate our actions with others in a precise manner and without much effort or thought (although of course, there are instances where more cognitive control is required from both parties). Joint action has been defined as social interaction between (at least) two individuals who coordinate their actions spatially and temporally in order to achieve a common goal (Sebanz et al., 2006), this can be verbal or non-verbal.

The underlying mechanisms of joint action encompass several key components (Sebanz et al., 2006). First mechanism is *joint attention*, which entails directing individuals' focus towards a common object or event of interest. Joint attention plays a pivotal role in establishing a mutual understanding of the object's shared representation. This facilitates the creation of a ‘perceptual common ground’ ensuring that both individuals perceive the same goal. Joint attention serves two primary functions within joint action. Firstly, it allows individuals to initiate the coordinated action, and secondly, maintain already established coordination of actions. Joint attention heavily relies on being able to infer what the other partner is attending to (Sebanz et al., 2006). Support for this mechanism comes from research displaying poorer joint action performance when joint attention to the common object of interest is limited (e.g., Clark & Krych, 2004). Further, another mechanism underlying joint action is action observation. Action observation occurs when a corresponding representation of the object or event of interest is created in the observer’s action system. In a joint action setting, this facilitates the comprehension of each other’s actions, creates ‘common ground’ in converging on mutual understanding of action goals, and aids the prediction of each other’s action outcomes. Specifically, ‘motor resonance’ – understanding the actions of others and their purpose in relation to a task – is crucial in predicting other’s actions. Empirical evidence has found that the gaze of the action observer precedes the action of their partner in for example a block-stacking task (Flanagan & Johansson, 2003). Finally, but possibly most important in joint action is *action adjustment.* Action adjustment allows to temporally and spatially adapt own actions in response to the partner’s actions and is reliant on feedback coming from actions of others. For example, Marsh et al. (2006) observed that a transition point of lifting planks of wood individually to in-pairs was dependent on the mean arm span of the *pairs* – that is pairs with a larger mean arm span moved to joint lifting later (longer planks) compared to those pairs with a smaller mean arm span. Joint action is influenced not just by an individual’s beliefs about their own abilities but also by the beliefs about what they can achieve in collaboration with others (Marsh et al., 2006).

The crucial aspects of predicting actions of others for successful joint action are about what, where, and when (Sebanz & Knoblich, 2009). The ‘what’ reflects predicting our partner’s type of action and the motivation behind it. The ‘when’ is important for coordination that entails precise timing (for example in a music band or orchestra). The ‘where’ is crucial when joint action is in a shared space to ensure that the shared movement tracks are more efficient and to prevent collisions (Sebanz & Knoblich, 2009).

The ‘when’ aspect of joint action (of particular interest for **Chapter 4** of the thesis) is all about taking into account the timing of the other person’s actions. This is related to *entrainment*, the idea that there is a tendency to *entrain actions* (synchronise or ‘fall into the same rhythm’), regardless of whether we are asked or not to do so with others (for example when rocking on a rocking chair, walking with the same foot when walking together, or even the synchronization of hanging clocks on the wall). The underlying mechanism of entrainment is the idea of coupled oscillators – “entrainment occurs as a consequence of direct, unmediated perception-action links between two or more systems that become coupled” (Sebanz & Knoblich, 2009, pp 359). Entrainment has the potential to foster stronger connections between individuals, thereby enhancing the coordination of intentional actions. However, entrainment has been seldom explored in an intentional joint action setting and the theories behind it are yet to be explored. This is also related to the hyperscanning technique, with studies finding synchronised or coherent brain-to-brain dynamics during social interactions. For example, Mu et al. (2016) reported significantly greater alpha phase locking value (PLV) (i.e., alpha phase coherence) between central electrodes of one of the participants and the posterior electrodes of their partner during the cooperative condition of button press synchronisation compared to the control individual task. Similarly, but with a verbal interaction paradigm, Kawasaki et al. (2013) observed enhanced synchrony in the amplitudes of theta and alpha activity during human-human compared to human-computer interactions. The concept of ‘falling into the same rhythm’ becomes notably apparent in the findings of Sänger et al. (2012), whose study revealed that brain-to-brain phase coherence was heightened when the requirement for guitar playing coordination was also increased.

1.3.1. Common coding and temporal expectations

Although entrainment can provide explanations for synchronised joint actions like for example motor finger tapping, other more discrete forms of joint action cannot be accounted for by this phenomenon (for example dancing together or carrying a heavy sofa up a flight of stairs). Instead, it has been proposed that every time we observe an action, we also generate a corresponding representation of this action. This is called 'common coding' (Sebanz & Knoblich, 2009). Previous research, both behavioural and neuroscience, has demonstrated that a common representation domain encodes both our own actions and the actions of others. The most famous evidence to support this claim comes from 'mirror neurons' in premotor cortex of monkeys, which become activated during own action performance as well as during the observation of other performing similar actions (Gallese et al., 1996). This is further supported by human studies, which also show that the observation of an action activates the action resonance network of inferior prefrontal, premotor, parietal and superior temporal cortex (e.g., Chao & Martin, 2000; Grafton et al., 1997), which are also responsible for motor execution (Grafton & Hamilton, 2007).

Beyond the capacity to generate action representations, the ability to form timed expectations about the outcomes of those actions is equally crucial (Davidson & Wolpert, 2003). This pertains to joint action in potentially two distinct ways. Firstly, it is plausible that we possess the capability to concurrently establish expectations for our own actions while accurately simulating and predicting the temporal dynamics of others' actions. Conversely, a more streamlined strategy for engaging in joint action involves directing attention to the collective outcomes of the collaborative performance (Knoblich and Jordan, 2003). Instead of imitating others' actions, their contributions are regarded as form of background 'noise'. Here joint action is achieved by concentrating on predicting the consequences of own actions and how these relate to the combined outcomes of the joint efforts (Sebanz & Knoblich, 2009). This is also in line with the Mutual predictive account that explains coherent hyperscanning results (Kingsbury et al., 2019). According to Sebanz and Knoblich (2009), joint learning is underpinned by the construction of joint performance models. Actions are often strategized based on the desired outcomes one aims to attain as opposed to more exact and immediate terms (such as the precise motor movements required). This revisits the fundamental essence of joint action, where the crux lies in the action coordination in time and space among individuals united by a shared objective. This thesis, specifically in **Chapter 4**, will examine how the emergence of common representation of the task goal within a joint action context is achieved, while also exploring the potential individual abilities required for better cooperative outcomes (**Chapter 5**).

1.3.2. Action adjustment in response to feedback

As discussed previously, one of the principles of joint action is action adjustment based on feedback. Empirical evidence is rather limited in unravelling the mechanisms that underlie joint action in response to feedback. However, a study by Knoblich and Jordan (2003) investigated the intricacies involved in joint action execution. Their experimental design tasked participants, both individually and in pairs, with keeping a target shape above a designated target line using speed adjustments – either slow down or speed up buttons. In the paired condition, the authors manipulated whether participants received feedback about their partner’s actions. The study found that the receipt of feedback concerning the actions of the other person was key to successful joint performance. Only with this feedback (alongside practice), pairs were able to match the performance success to the performance achieved as individual. This empirical evidence gives rise to the notion that successful joint action requires the linkage of consequences of our own and others’ actions.

The neurobiological underpinnings of feedback in response to own actions have been studied in ‘one-brain’ studies extensively. EEG changes frequently associated with feedback processing include the feedback-related negativity (FRN) and the midline-theta response. The FRN is an event-related potential that occurs in response to feedback. It is typically maximal over frontal-central electrodes and has a maximal peak around 250 ms after the onset of the feedback. This response is prominent for negative feedback and is thought to reflect the brain's evaluation of errors or the violation of expectations (Holroyd & Coles, 2002; Talmi et al., 2013). Several studies have incorporated a task that requires participants to coordinate joint action outcomes, but only examined the neural mechanisms of one of the participants. Van Schie et al. (2004) found that the FRN component occurred in participants that observed errors made by others. The FRN component is also present in cooperative contexts following a partner’s error (Czeszumski et al., 2019; Itagaki & Katayama, 2008; Picton et al., 2012), although reduced compared to own errors (Loehr et al., 2015) and not sufficient to lead to action adjustment (Picton et al., 2012). However, due to its relatively small amplitude and variability across trials, accurately capturing the FRN waveform and its characteristics becomes difficult on a trial-to-trial basis.

The onset of the feedback cue also induces oscillatory activity (i.e., time-locked, but not necessarily phase-locked) in the 3-7 Hz (theta range), following the same time-course of the FRN, and with a maximal midline frontal distribution. While the feedback induced midline-theta overlaps with the FRN in time and space it has been suggested to reflect the initiation of behaviour adjustments (i.e. cognitive control) overriding ‘status quo’ responses (Cavanagh et al., 2013; Cavanagh & Frank, 2014). For example, van de Vijver et al. (2011) showed an increased theta power after negative feedback (compared to positive feedback), indicating a possible requirement for behavioural adjustment. This is further supported by Cohen (2016) who proposed the involvement of mid-frontal theta in scenarios necessitating the engagement of the monitoring system for error correction.

The investigation into the brain-to-brain neural mechanisms that underlie joint action, specifically implementing external feedback, remains constrained. To my knowledge, only two studies employed feedback in an EEG hyperscanning joint action setting. Firstly, Mu et al. (2016) examined inter-brain synchrony during a coordination task while implementing feedback. However, their findings centred predominantly around the coordination aspect, failing to probe the nuanced brain-to-brain dynamics prompted by feedback. Secondly, Balconi et al. (2018) examined the effect of external feedback on synchronised joint action and the related EEG inter-brain dynamics. The external feedback (compared to no feedback) decreased inter-brain power connectivity in the theta and delta bands in the frontal regions. However, in this study the generated feedback was not based on performance but rather superficially created. The existing lack of comprehensive studies examining the neural underpinnings of the interplay of joint action and feedback becomes apparent. Examining this topic does not only deepen our understanding of the underlying mechanisms at play but also holds significant implications for refining cooperative processes and enhancing communication. Thus **Chapter 3** will examine the underlying brain-to-brain dynamics of developing convergence on task goals based on external feedback.

1.3.3. Individual differences in joint action

Existing literature has identified various individual differences that may drive the success of joint action as well as social interaction outcomes, for example personality traits such as agreeableness and honesty-humility (Thielmann et al., 2014) or general reciprocity (Salazar et al., 2022). Given that Theory of Mind (ToM) and joint action share similar underlying mechanisms, understanding and predicting of own and other actions, it is not hasty to assume a relationship between them. This question has previously been posed for future research in a seminal paper by Sebanz et al. (2006), if and if so, how, does joint action rely on Theory of Mind. This is tested and discussed in **Chapter 4** of this thesis.

ToM has been defined as the ability to assign emotions and mental states, including beliefs and intentions, to others (Gallese & Sinigaglia, 2011). There is a longstanding tradition of devising tasks to measure ToM abilities in children, for example the False beliefs task (Wimmer & Perner, 1983). However, within healthy young adults, these tasks often yield ceiling effects leading to a limited range of abilities, which hinders their applicability to other cognitive domains. In the last decade, a shift has transpired towards more intricate and nuanced assessments capable of capturing the variance in ToM abilities amongst adults. For example, the animated triangles task (Castelli et al., 2000) requires participants to expound upon the intentions behind the movements of geometrical shapes. The Reading the Mind in the Eyes (RMET; Baron-Cohen et al., 2001) tasks individuals with matching emotional verbal descriptors to images of eyes. Further, Dziobek et al. (2006) has developed an audio-visual task called the Movie Assessment for Social Cognition (MASC), in which participants watch short clips of social interactions and are tested on the intentions, feelings and beliefs of the characters. Distinct from static depictions, the MASC challenges the judgment of dynamic social scenarios complete with body language, expressions, and conduct. For this reason, we decided to measure ToM abilities using the MASC as it encapsulates the intentions and beliefs parallel to genuine life scenarios. This dedication to real-world dynamics was equally paramount in the formulation of our innovative 2-player verbal cooperation game, where participants collaboratively integrated separate information fragments to achieve cooperative success. Again, this deliberate design aimed to mirror genuine cooperative scenarios, featuring measurable cooperative facets.

Previous research has shown a clear link between ToM and cooperation in children, with cooperation being measured in Ultimatum games (Takagishi et al., 2010) or spatial rotation tasks (Grueneisen et al., 2015). The picture is less clear in adults due to the difficulty of reliably measuring ToM. Nevertheless, some studies have reported that enhanced mindreading abilities are linked to better social cooperation (Paal & Bereczkei, 2007). Although, existing literature thus hints at a relation between ToM and cooperation, the precise nature and degree to which ToM propels cooperative behaviour remains vague in healthy adults. Further, earlier investigations have predominantly focused on ToM differences of *individuals*. Given the inherent collaborative nature of social interactions, it is crucial to consider the collective profiles of ToM capabilities within the pairs. Thus, **Chapter 4** aims to directly investigate the relationship between ToM and verbal cooperation via a newly developed verbal cooperation paradigm. This study aims to advance the trajectory of social interaction research by aligning the verbal cooperation measure with authentic real-life cooperative dynamics. This is achieved by exploring how the ToM abilities pairings contribute to the cooperative success outcome.

1.4 Thesis summary

In summary, the central objective of this thesis revolves around the investigation of language, communication and social interaction. Firstly, I investigate the age-related oscillatory differences in language comprehension, with a specific emphasis on lexical retrieval and semantic binding processes (**Chapter 2**). Secondly, I investigate the brain-to-brain neural dynamics that forecast the outcomes of joint action/ cooperation (**Chapter 3**). Lastly, I aim to identify the distinctive individual attributes, specifically Theory of Mind, that contribute to successful cooperation and joint action (**Chapter 4**Anticipating the key findings, we found age-related and individual specific impacts on different communication facets and linguistic processes, alongside key signatures of neural dynamics between individuals that forecast the success of cooperative joint action, as well as key contributions of ToM to joint action outcomes.

CHAPTER 2: HOW THE HEALTHY AGEING BRAIN SUPPORTS SEMANTIC BINDING DURING LANGUAGE COMPREHENSION

Semantic binding refers to constructing complex meaning based on elementary building blocks. Using EEG, we investigated the age-related changes in modulations of oscillatory brain activity supporting lexical retrieval and semantic binding. Young and older adult participants were visually presented two-word phrases, which for the first word revealed a lexical retrieval signature (e.g. *swift* vs. *swrfeq*) and for the second word revealed a semantic binding signature (e.g. *horse* in a semantic binding “swift *horse*” vs. no binding “swrfeq *horse*” context). The oscillatory brain activity associated with lexical retrieval as well as semantic binding significantly differed between healthy older and young adults. Specifically for lexical retrieval, we found that different age groups exhibited opposite patterns of theta and alpha modulation, which as a combined picture suggest that lexical retrieval is associated with different and delayed signatures in older compared to young adults. For semantic binding, in young adults we found a signature in the low-beta range centred around the target word onset (i.e. a smaller low-beta *in*crease for binding relative to no binding), while in healthy older adults we found an opposite binding signature about ~500ms later in the low- and high-beta range (i.e. a smaller low- and high-beta *de*crease for binding relative to no binding). The novel finding of a different and delayed oscillatory signature for semantic binding in healthy older adults reflects that the integration of word meaning into the semantic context takes longer and relies on different mechanisms in healthy older compared to young adults.

Published in:

Markiewicz, R., Segaert, K., & Mazaheri, A. (2021). How the healthy ageing brain supports semantic binding during language comprehension. European Journal of Neuroscience, 54(11), 7899–7917. https://doi.org/10.1111/ejn.15525

Data availability:

Stimuli and data are available here: https://osf.io/f8grv/

2.1 Introduction

Healthy ageing is accompanied by decline across a number of cognitive domains, such as your memory for events and the speed with which you process information (Salthouse, 1996; Waters & Caplan, 2005). Language is a crucial aspect of cognition, but the picture of how ageing affects language is a complex one. Older adults get better at some aspects of language, such as knowing more words (Brysbaert et al., 2016), while other skills clearly deteriorate, for example, accessing all the word-related information you need for production (Hardy et al., 2020; Segaert, et al., 2018). At the same time, many other language abilities, including sentence comprehension, appear relatively unchanged by healthy ageing (Peelle, 2019; Shafto & Tyler, 2014). For example, sentence comprehension performance has been demonstrated to be comparable between older and young participants, unless the stimuli are presented at a rapid rate (Tun, 1998; Wingfield et al., 2003) or with background noise (Tun, 1998). The complex behavioural picture for language function is difficult to reconcile with the widespread structural decline in language-relevant brain regions (Antonenko et al., 2013). Even when language performance appears unchanged in older adults, it is likely supported by different functional neural processes from those in young adults (Peelle, 2019). The aim of the current electroencephalography (EEG) study is to investigate the differences between healthy older and young adults in the neural processes involved in semantic comprehension.

When we combine words, the meaning of an individual word (e.g. flat) can be altered by the meaning of a following word (e.g. flat tire vs. flat note) such that the combined meaning is more than the mere sum of its parts (Hagoort et al., 2009; Keenan, 1979). This illustrates the unique and expressive power of language: we have the ability to combine words in novel ways to create sentences. In other words, language users construct complex meaning from more elementary semantic building blocks (Hagoort, 2020; Hagoort et al., 2009). This ability forms the basis for communication and social interactions. Understanding the meaning of a multi-word utterance requires a process we refer to here as semantic binding. Lexical retrieval of information (including semantic, syntactic, and phonological details) from long term memory is required. The lexically retrieved information about single words needs to be integrated into a representation of a multi-word utterance. This process has also been referred to as merge (Chomsky, 1995; Zaccarella & Friederici, 2015) or unification (Hagoort, 2005). ERP research has demonstrated that word meaning is assembled into larger meaning representations in less than 500ms (Kutas & Hillyard, 1980), with this process immediately taking into account information from a wide range of sources, including world knowledge and discourse (Hagoort et al., 2009).

Combining words together also requires the consideration of syntactic information, including tense, aspect and agreement (Segaert, et al., 2018). Therefore, semantic binding cannot exist in the absence of syntactic binding and disentangling the two is difficult. Previous literature has disagreed on the best solution to unravel semantic from syntactic binding and therefore assigning semantic/syntactic binding-specific processes to observable effects has been difficult (Bemis & Pylkkänen, 2011). The present study examines the linguistic composition involved in adjective-noun minimal phrases. Although we refer to the observed effects as semantic binding effects, it is important to note that semantic and syntactic binding are conflated within the phrases and both processes are somewhat either simultaneously present or not.

Neuroimaging studies employing fMRI have been able to provide a wealth of information about the location of brain areas likely associated with semantic binding in young adults. Previous investigations have found evidence that semantic binding requires the exchange and integration of information in a large network of frontal and posterior areas, including left inferior frontal gyrus, bilateral superior and middle temporal gyri, anterior temporal lobe and angular gyri (Baggio & Hagoort, 2011; Lyu et al., 2019; Menenti et al., 2011; Pylkkanen, 2019; Tyler & Marslen-Wilson, 2008). The functional neural characteristics supporting specific language functions in healthy older adults, have often been found to differ from those in young adults (Antonenko et al., 2013; Peelle, 2019; Shafto & Tyler, 2014; Tyler et al., 2010; Wingfield & Grossman, 2006). With age, structural changes occur in language-relevant brain regions. In the context of these structural changes, it would be unlikely that successful performance in older adults is achieved with identical neural processes as in young adults (Peelle, 2019). Generally, the literature shows a more widespread pattern of activity in healthy older adults relative to young adults (e.g., Cabeza et al., 2002; Davis et al., 2008). Different views exist on how to interpret these age-related changes in brain activity: the appearance of more diffuse activity in older adults may reflect a general decline in neural efficiency (i.e. dedifferentiation), alternatively (though not mutually exclusive) increased engagement of brain regions may reflect focused recruitment as a means to compensate for neurocognitive decline (i.e. compensation) (Wingfield & Grossman, 2006).

One limitation of fMRI is the slow time course of the hemodynamic response (1.5-5 seconds) which limits what information it can provide about ‘when’ the specific neural processes involved in semantic processing are occurring. While EEG as a neuroimaging tool does not have the spatial resolution of fMRI, it does provide a real-time window into the neural activity underlying cognition. Previous EEG studies investigating how the brain supports semantic comprehension, have primarily looked at event-related brain potentials (ERPs) which represent brain activity phase-locked to the onset of words. As briefly mentioned above, these studies have consistently found that word meaning in young adults is integrated into the meaning of a larger multi-word utterance at around 400-500ms after the relevant word, as indexed by the N400 ERP (Kutas, Hillyard, 1980; Kutas & Federmeier, 2011). Previous studies have also elucidated several relevant aspects of how older adults comprehend sentence-level meaning. Healthy older adults do extract and make use of contextual semantic information (Stine-Morrow et al., 1999), but there are differences (compared to young adults) with respect to when and how this happens. Sentential context manipulations (i.e. the strength of contextual constraint for sentence-final words) elicit reduced and delayed N400 effects for older (compared to young) adults (Federmeier & Kutas, 2005; Wlotko & Federmeier, 2012). Moreover, effects of message-level congruity on the N400 are delayed by over 200 ms in older adults (Federmeier et al., 2003). Ageing furthermore affects processing of compositional concreteness, i.e. processing of the second noun in a noun-noun pair, in function of whether the first was concrete versus abstract (e.g. alias-battle vs. skate-battle) (Lucas et al., 2019), further suggesting that there are age-related changes in compositional semantics in healthy older (compared to young) adults.

There may be multiple (not mutually exclusive) sources of the observed age-related changes in how the brain supports making use of contextual semantic information. Older adults may engage different functional neural processes to support semantic binding and maintain a message-level meaning representation while processing incoming information. In addition, older adults may be less able to use prediction mechanisms during language comprehension. Several studies have provided support for the latter (Federmeier et al., 2002; Wlotko et al., 2012). In the present study, we focus on the former: do healthy older and young adults engage different neural mechanisms for semantic binding?

We will answer this question by investigating oscillatory activity. The EEG signal contains oscillatory activity (i.e., rhythms) which are hypothesized to play a vital role in how the brain carries out cognition (Mazaheri, et al., 2018; Siegel et al., 2012). Investigating the oscillatory (i.e., spectral) changes in the EEG allows for capturing activity that is time-locked but not necessarily phase-locked to experimental events (i.e., the onset of words). Studies focusing on the spectral changes in the EEG have found that the exchange and integration of information required for semantic binding, involves modulations in the oscillatory power in the theta (~4-7 Hz), alpha (~8-12 Hz), and beta (~15-30 Hz) bands (for comprehensive reviews: Meyer, 2018; Prystauka & Lewis, 2019; for a detailed overview see Weiss & Mueller, 2012). Modulations of each frequency band are thought to reflect different language comprehension related processes. However, it is not always easy to map specific roles onto definitive oscillatory ranges. The theta frequency range is related to memory retrieval and processing demands, whereas alpha parallels attentional processes and storage of phrases (Meyer, 2018). Previous studies have also further subdivided the beta frequency ranges into more narrow bands: ~13-20Hz (i.e., low beta), and 20-30Hz (i.e., high beta) (Poulisse et al., 2020; Segaert, Mazaheri, et al., 2018; see Weiss & Mueller, 2012 for a detailed overview). Low beta is related to higher-order processing. In language comprehension this translates to linking past and present input (i.e. binding including semantic features (Bastiaansen & Hagoort, 2006; von Stein et al., 1999; Weiss & Mueller, 2003) and syntactic unification (Bastiaansen et al., 2010)). High beta has previously been related to processing of action/motor-related language (Elk et al., 2010).

Although there are a number of previous studies that reveal the oscillatory signatures of semantic binding in young adults, the pattern across studies is not always clear cut. Firstly, several previous studies (but not all, e.g., Kielar et al., 2018; Wang et al., 2018) have shown that increased theta power is associated with semantic anomalies (Bastiaansen & Hagoort, 2015; D. J. Davidson & Indefrey, 2007; Hagoort et al., 2004; Hald et al., 2006; Wang, Zhu, et al., 2012). This is thought to reflect the increased effort (i.e. neural resources) required to integrate semantically incongruous items into the wider context. Furthermore, increased theta power has also been linked with syntactic violations (e.g., Bastiaansen et al., 2002; Kielar et al., 2015; Lewis et al., 2016), supporting the notion that this theta signature may reflect a more general violation detection mechanism (Prystauka & Lewis, 2019). Furthermore, the importance of beta oscillations in linguistic composition processing has previously been highlighted (Lewis et al., 2015; Lewis & Bastiaansen, 2015). Lewis and Bastiaansen (2015) proposed the hypothesis that power changes in the beta band are linked to linguistic information maintenance. More specifically, active maintenance of linguistic cues leads to beta power increases. The language processing system actively maintains the current cognitive state due to the greater processing demands arising from more complex clauses. This notion is supported by a study by Meyer et al. (2013), who have found an increase in the beta power (13 to 20Hz) at the point of the verb in sentences with long-distance dependencies (between argument and verb) compared to sentences that contained short-distance dependencies. Long-distance dependency clauses are computationally more demanding and complex (compared to short-distance sentences), directing the language comprehension system towards active maintenance of the processing mode.

On the other hand, any change of linguistic input is related to beta power decreases. Lewis and Bastiaansen (2015) suggested that semantic (and syntactic) violations are clear cues to the system indicating a need for change. Therefore linguistic violations lead to decreases in beta power compared to instances where semantic (or syntactic) violations are not present. This theory is supported by multiple empirical studies. For example Kielar et al. (2014) (see also Kielar et al., 2015, 2018) investigated the effect of violations on oscillatory responses using semantically correct (e.g. “A new computer will *last* for many years”) vs. semantically anomalous (e.g. “A new computer will *paint* for many years”) sentences. They found that the semantic violations elicited power decreases in the 8 to 30Hz range and were maximal ~500-1000 ms post the target word onset over parietal sites. Furthermore, Luo et al. (2010) provided additional support for the notion that semantic violations lead to low beta (16 to 20Hz) decreases immediately after (0-200 ms) the target word onset as well as in a later window (~500 ms later) using Chinese semantically congruent vs. incongruent sentences. This beta modulation in the later time window (around the N400 effect) has also been found by Wang, Jensen et al. (2012) with anomalous words eliciting a decrease in the beta power. On the other hand, Lam et al. (2016) reported an opposite result, showing stronger beta power decrease for real sentences compared to word lists with the most prominent effect at ~350ms.

The authors related this effect to stronger neural activation for sentences (compared to word lists) reflecting the unification of semantics and syntax. Lastly, several previous studies suggested that the beta signature for maintenance and binding of linguistic information extends into the alpha power range (Gastaldon et al., 2020; Kielar et al., 2014; Lam et al., 2016; Luo et al., 2010; Katrien Segaert, Mazaheri, et al., 2018).

In the present study, we aim to investigate modulations in oscillatory brain activity (with pre-defined frequency bands based on previous literature: theta (4-7Hz), alpha (8-14Hz), low beta (15-20Hz), and high beta (20-25Hz)) during semantic binding in healthy older compared to young adults, using a minimal two-word phrase paradigm. We present target words (e.g. horse) in a semantic binding (e.g. swift horse) vs. no semantic binding context (e.g. swrfeq horse). In both cases, for the target word, retrieval of lexico-semantic information from memory takes place. However, only in the binding condition a complex meaning representation can be built for the phrase, based on the elementary building blocks of each individual word. It is important to note here again that syntactic binding is also present in the semantic binding context as the two are not easily disentangled. As the semantic content, and not the syntactic features of the phrases, is manipulated in the present paradigm, we refer to the conditions as semantic versus no semantic binding. Within the semantic binding condition we also manipulate whether the phrase is plausible. Although secondary, the use of this paradigm also allows an investigation of lexical retrieval effects and the recognition of the word form (e.g. swift vs. swrfeq). So we will report these findings as well.

The computation for a two-word phrase forms the foundation of binding in the context of increasing complexity. Investigating elementary semantic binding by means of a minimal phrase paradigm offers the advantage of focusing on the binding process while minimizing contributions of other processes involved in sentence comprehension, such as working memory load and the ability to use predictions. This advantage is particularly salient when investigating age-related changes in how the brain supports online sentence comprehension, given that working memory and the ability to use predictions are also impacted by age. Bemis & Pylkkänen (2011) conducted one of the first studies with a minimal paradigm, and compared nouns in a minimal binding context (e.g. red boat) versus a wordlist condition (e.g. cup boat). This inspired many other studies to use similar designs (Bemis & Pylkkänen, 2013; Pylkkänen et al., 2014; Segaert, et al., 2018; Zaccarella et al., 2017; Zaccarella & Friederici, 2015). Poulisse et al., (2019, 2020) used this approach to investigate how healthy ageing impacts minimal syntactic binding.

In line with previous literature on semantic comprehension in young adults (Lewis et al., 2015), we expected to see a greater beta power decrease (particularly in the low-beta (15-20Hz) frequency band) in the no semantic binding compared to the semantic binding condition. We made no predictions concerning a modulation in the theta range, since our paradigm does not manipulate violations (Prystauka & Lewis, 2019), but rather, successful binding versus no binding (with previous behavioural performance results demonstrating that no binding occurs for pairing a pseudoword with a real word (Poulisse et al., 2019; Katrien Segaert, Mazaheri, et al., 2018)). If previously observed age-related changes in making use of contextual semantic information are at least in part due to a change in how the brain supports semantic binding, then we expect to see different oscillatory signatures for healthy older versus young adults, in the semantic binding versus no semantic binding conditions, however this aspect of our study is exploratory and entirely novel, making it difficult to make concrete predictions about the direction of power changes in the alpha or beta range (Beese et al., 2019; Poulisse et al., 2020).

2.2 Materials and methods

2.2.1 Participants

Thirty-three young adults and 32 healthy older adults took part in the study. However, seven participants were excluded from the analysis due to: (a) excessive EEG artefacts in the recordings (N=4), and (b) being bilingual (N=3). The participants included in the analyses were 29 young adults (2 males, aged 18-24) and 29 healthy older adults (13 males, aged 63-84) (see Table 2.1 for more information). All participants were right-handed, British-English monolingual speakers with normal-to-corrected vision and no neurological or language impairments. All older adults scored above 26 out of 30 in the Montreal Cognitive Assessment (MoCA) test (M = 27.79, SD = 1.01, min = 27, max = 30) (Note that we have adopted a conservative threshold, as recommended in prior research, where scores ≤ 26 indicate a risk of mild cognitive impairment or dementia (Smith et al., 2007). However, it is worth mentioning that some studies propose an even more lenient threshold of 23 (Carson et al., 2018)) .

Table 2.1 Demographic and cognitive characteristics for young and older adult participants

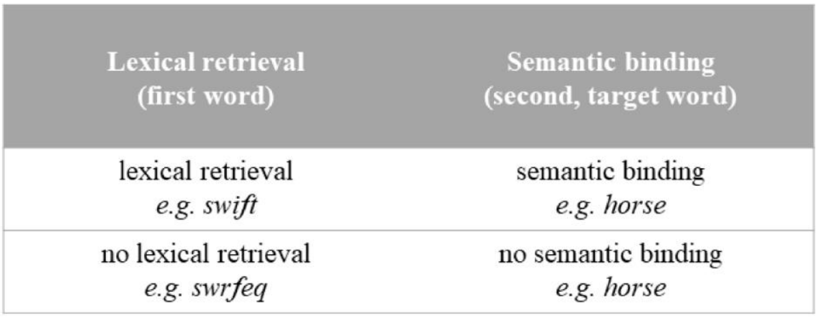
	Young adults (N=29)		Older adults (N=29)			
	M	*SD*	*M*	*SD*	*t-value*	*P*
Age (years)	19.5	1.5	73.6	5.8		
Years of education	14.52	1.5	15.45	3.08	-1.46	*.15*
Processing speed	81.48	11.57	65	13.17	4.97	< *.001*
NART	25.81	5.07	37.05	4.61	-8.82	< *.001*
Working memory	4.42	.66	4.97	.92	-2.62	<*.05*

The young adults were Undergraduate students from the University of Birmingham and took part in the study for course credits. The older adults were from the Patient and Lifespan Cognition Database and were compensated for their time with cash payments. Participants signed informed consent, which followed the guidelines of the British Psychology Society code of ethics, and the experiment was approved by the Science, Technology, Engineering, and Mathematics (STEM) Ethical Review Committee for the University of Birmingham (Ethics Approval Number: ERN_15-0866).

There was no significant difference in the number of years spent in education between the younger and the older adults. In line with expectations, young adults outperformed older adults in processing speed (Weschler Adult Intelligence Scale-IV processing speed index), whereas older adults outperformed young adults on the National Adult Reading Test (Nelson, 1982). Surprisingly, older adults also outperformed young adults in the working memory tasks (i.e. the average combined score of the backward digit span and subtract 2 span tests) (Waters & Caplan, 2003), which could be attributed to young adults being less motivated when they were participating in the tasks (in line with similar findings reported previously: Heyselaar et al., (2020)).

2.2.2 Design, materials and task

We created a minimal language comprehension paradigm with two-word-phrases. Each phrase included two words, where the target word was always the second word. The design of the study, with example stimuli, is illustrated in Figure 2.1. We manipulated lexical retrieval (comparing real words to letter strings) and semantic binding (comparing the target word in a semantic binding context to a no semantic binding context). We were primarily interested in the effects of semantic binding, but since our two-word phrase paradigm allows examining the effects of lexical retrieval and the recognition of the word form also, we report these effects below as well. Within the semantic binding condition, we furthermore manipulated whether semantic binding was plausible (e.g. swift horse) or implausible (e.g. barking horse).



Lexical retrieval (first word)	**Semantic binding (second, target word)**
lexical retrieval *e.g. swift*	semantic binding *e.g. horse*
no lexical retrieval *e.g. swrfeq*	no semantic binding *e.g. horse*

Figure 2.1 Example word-pairs in each condition.

To ensure participants paid attention to the word-pair stimuli throughout the experiment, we included questions about the word-pairs on a subset of the trials (22% of all trials). The questions asked “Did you just see [word pair]”. There were no significant differences between young and older adults in response accuracy (young adults: mean = 94.83, SD = 0.5; older adults: mean = 96.48, SD = 0.2; $t(42.109) = -1.557$, $p = .127$) or reaction times (young adults: mean = 1718.78, SD = 503.64; older adults: mean = 1851.45, SD= 488.92; $t(56) = -1.018$, $p = .313$). All the participants scored higher than 80%. From this we can conclude that young and older adults paid close attention to the language stimuli as they were being presented to them throughout the experiment.

We verified our plausibility manipulation in an online rating study with 57 respondents. The online survey asked to rate the plausibility of the two-word phrases, where 1 = ‘Completely implausible’, 2 = ‘Somewhat implausible’, 3 = ‘Somewhere in between’, 4 = ‘Somewhat plausible’, and 5 = ‘Completely plausible’. The plausibility ratings were significantly different for the plausible (M=3.87, SD=.29) and the implausible word-phrases (M=1.93, SD=.66); Welch’s F(1, 101.65) = 526.36, $p < .001$.

In each condition, about half of the target words were animate, the other target words were inanimate. The exact trial distribution was as follows: inanimate-plausible (N=45), inanimate-implausible (N=46), inanimate-letter string (N=44), animate-plausible (N=42), animate-implausible (N=47), and animate-letter string (N=46). The list of plausible and implausible adjectives was matched for word frequency using the CELEX database (Baayen et al., 1993) (plausible *mean* = 28.16, *SD* = 40.63, implausible *mean* = 27.87, *SD* = 36.67), number of syllables (plausible *mean* = 1.74, *SD* = 0.49, implausible *mean* = 1.79, *SD* = 0.52) and number of letters (plausible *mean* = 5.87, *SD* = 0.79, implausible *mean* = 5.64, *SD* = 0.71).

Three versions of the experiment were created, where the same word-pairs were presented in different orders. Participants were randomly assigned to one of the three versions. The paradigm intended to present 60 attention-questions for each of the versions of the experiment. However, due to an error in creating the question lists, the number of questions differed slightly per version (either 61 or 62 questions). The questions were not used in any of the EEG analyses. A full stimulus list, each of the 3 versions of the experiment, with the exact attention questions asked, can be downloaded from https://osf.io/f8grv/.

2.2.3 Procedure and trial timing

We presented our experiment using E-prime 2.0. Figure 2.2 depicts the duration of time that each element of the trial was presented for (top row) and the presentation time of each element in the trial when the EEG epoch was locked to the onset of word 1 (bottom row). The task consisted of 270 trials divided into 9 blocks. In between each block, we offered the participants a break.

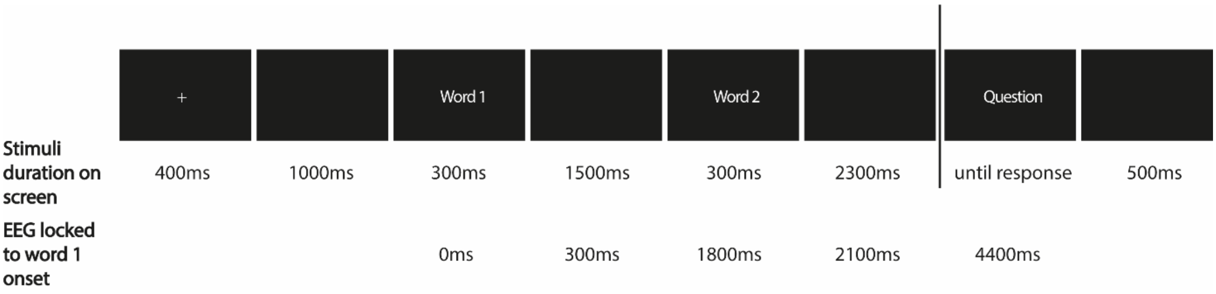


Figure 2.2 Trial presentation of the minimal two-word phrase paradigm. The questions appeared in 22% of the trials. The top row ("Stimuli duration on screen") depicts the on-screen time duration (in ms) of each trial element. The bottom row ("EEG locked to word 1 onset") depicts the presentation time (in ms) of each trial element when the trial/ EEG epoch is locked to the onset of word 1.

Upon the start of the experimental session, participants were fitted with a 64-electrode EEG cap. Once the EEG set-up was finished participants sat in a sound proof booth 70 cm from the monitor where the computerised task took place. Participants were instructed to read in silence word-pairs (e.g., swift horse) appearing on the screen. They were told that from time-to-time they would see a question on the screen regarding the word-pair that they had just seen (e.g. Did you just read 'swift horse?'). Participants were able to indicate 'yes' or 'no' using a button box. Participants completed a practice block first to familiarise themselves with the paradigm (30 word-pairs, with 9 questions, which were similar in nature but different from the experimental stimuli), which was followed by the actual experiment. Following the computer task, participants completed Working Memory tests (i.e. the Backward digit span task and the subtract 2-digit span task) (Waters & Caplan, 2003), National Adult Reading Test (NART) (Nelson, 1982), and the Weschler Adult Intelligence Scale-IV processing speed index (Weschler, 2008). In addition, the older participants also completed the Montreal Cognitive Assessment (MoCA) (Nasreddine et al., 2005).

2.3 EEG recording

EEG was recorded using Waveguard caps containing 64 cap-mounted Ag/AgCI electrodes (10-20 layout, including left and right mastoids). Horizontal eye movements were measured by two electrodes placed on the outer left and right canthi. Vertical eye movements were recorded by two electrodes placed above and below right eye. The EEG recording was acquired with online reference to the CPz channel. The signal was amplified with the ANTneuro EEGosports amplifier system and recorded using EEGo software (Advanced Neuro Technology). The signal was obtained at a sampling rate of 500Hz, with a 30Hz low-pass filter (24 dB/octave) and a 0.05Hz high-pass filter, implemented in the EEGosports firmware. We aimed to keep the impedances below 10 kΩ.

2.4 EEG analysis

2.4.1 EEG pre-processing

The EEG pre-processing was performed using EEGLAB 14.1.2b (Delorme & Makeig, 2004) and Fieldtrip toolbox 2018-07-16 (Oostenveld et al., 2011). The data were epoched to the onset of the first word (-1.5sec to 4.4sec) and later offline re-referenced to the average of all of the channels, where the mastoid and bipolar electrodes were excluded from the re-referencing. EEGLAB was used for manual inspection and rejection of trials with non-physiological artefacts. The average number of removed trials was 26.34 (SD=21.93) per participant due to artifacts. Ocular artifacts were removed based on the scalp distribution using independent component analysis ("runica") in EEGLAB. The average number of removed components was 2.07 (SD=.81) for each participant. Channels TP8 and TP7 were removed before completing any analyses due to poor or no signal from these channels across the participants.

2.4.2 Time-frequency representations of power

Time-frequency representations (TFRs) of power were performed using the FieldTrip ‘mtmconvol’ method with a sliding time window. The Hanning taper was applied to the adaptive time window of 3 cycles per each frequency of interest (i.e. the length of the window at each frequency of interest is equal to 3/f s) for every trial. Similar approaches were used previously by Mazaheri et al. (2009), Poulisse et al. (2020), and van Diepen et al. (2015). The analysis included the frequency of interest of 2Hz to 30Hz in steps of 1Hz, and the time of interest of -1.5 to 4 sec in steps of .05 sec. We calculated the changes in oscillatory power locked to the onset of the stimulus (i.e. word one) in relation to the change in power from baseline. The data were baseline corrected to a window of -600 to -100 ms prior to stimulus onset (i.e. presentation of first word). This was predefined and thus applied within conditions and age groups.

We statistically examined baseline differences between groups (collapsed across conditions) using non-parametric cluster-based permutation tests (see below for further explanation) prior to correcting the baseline window. We used an averaged a priori time window of -600 ms to 0 sec (where 0 was the onset of word 1) within the following pre-defined frequency bands: theta (4-7Hz), alpha (8-14Hz), low beta (15-20Hz), and high beta (20-25Hz) in the cluster-based permutation tests. The cluster-based permutation tests did not reveal any significant differences between groups (healthy older adults vs. young adults collapsed across all conditions) in the averaged baseline window (-600 ms to 0).

To ensure that the observed oscillatory changes were not just the spectral representation of the ERPs, the ERP components were subtracted from the TFR (Mazaheri & Picton, 2005). The subtraction was achieved by first generating the time frequency decomposition of the ERP data for each condition and participant separately. Next, the time frequency power spectra of the ERP was subtracted from the time frequency power spectra of the EEG signal for each condition. The subsequent power changes in the time-frequency domain were used to generate time frequency power spectra differences between experimental conditions (lexical retrieval/semantic binding vs no lexical retrieval/ semantic binding; plausible semantic binding vs implausible semantic binding) for each group separately.

Finally, the statistical differences of the experimental condition differences in the power changes in the time-frequency domain were assessed by using a non-parametric cluster-based permutation test (using FieldTrip toolbox) (Maris & Oostenveld, 2007). Each channel/time/frequency pair locked to the onset of the first word for the difference between each experimental condition (i.e. binding vs no binding) was compared using an independent (two-tailed) samples t-test (for young vs older adults) with a threshold at 5% significance level. Significant pairs were then clustered (cluster was defined based on proximity in space using the triangulation method i.e., having at least two significant electrodes that were adjacent to each other) and participants' labels of each cluster were randomly shuffled using 1000 partitions. The Monte Carlo P values were calculated using the highest sum of the test statistic. An equivalent dependent (two-tailed) samples t-test was used to compare the experimental manipulations within each group separately in order to extract between condition effects. The time window used to assess the statistical differences in time-frequency power for (1) lexical and semantic binding manipulation was 0 (onset of word 1) to 3.2 sec, where onset of word 2 occurred at 1.8 sec, and (2) plausibility manipulation was 1.8 sec (onset of word 2) to 3.2 sec. The above analysis was performed within the following pre-defined frequency bands: theta (4-7Hz), alpha (8-14Hz), low beta (15-20Hz), and high beta (20-25Hz) consistent with previous studies (Poulisse et al., 2020; Segaert, et al., 2018). The average number of trials across all participants included in the final analysis for the lexical retrieval/ semantic binding condition was 162.87 (SD = 15.27), for the no lexical retrieval/ no semantic binding condition 80.79 (SD = 7.4), for the plausible condition 79 (SD = 7.6), and for the implausible condition 83.86 (SD = 8.05). Note here that we checked whether any of the observed condition effects in the data are not spurious differences attributed to the imbalance of signal-to-noise ratio (SNR) (due to the imbalance of the number of trials in lexical retrieval/semantic binding and no lexical retrieval/ no semantic binding). To do this, we used a random re-sampling approach to equate the trials in both conditions. We then completed between condition (lexical vs. non lexical) dependent non-parametric cluster-based permutation tests within each group separately and found similar condition effects in each group (see supplementary materials for the full report of these results, and Suppl. Figure SM.2) as those reported here. Therefore, we can be confident that any observed condition differences in each group are not due to the imbalance of SNR.

2.5 Results

We first visually inspected the TFRs and qualitatively describe the power modulations. The onset of word 1 and word 2 generated an increase in theta (4-7Hz) power, followed by a suppression of alpha power (8-14Hz), irrespective of condition in both age groups (Figure 2.3). Consistent with previous studies (e.g., Bastiaansen et al., 2005, 2008; Hermes et al., 2014; Mazaheri, et al., 2018), the theta power increase peaked at around 0.2 sec post word onset and was maximal over the occipital channels. Also in line with previous work, the alpha power suppression peaked at around 0.5 sec post word onset and was maximal over the occipital channels (Davidson & Indefrey, 2007; Mazaheri, et al., 2018).

Our analysis approach was as follows. We focused our analysis on the oscillatory changes in the EEG associated with lexical processing (e.g., 1st word: swift vs. swrfeq) and semantic binding (e.g., horse when preceded by swift vs. when preceded by swrfeq). To help with the interpretation of the between-group effects, in the main text we only describe the significant oscillatory differences between conditions within groups (Figure 2.3A and B) *if* they were also significantly different between young and older adults (Figure 2.3C). A comprehensive and more detailed description of condition differences within each age group separately can be found in the supplementary materials (and Suppl. Figure SM.1).

In what follows, we first describe the lexical retrieval results, and then the semantic binding results.

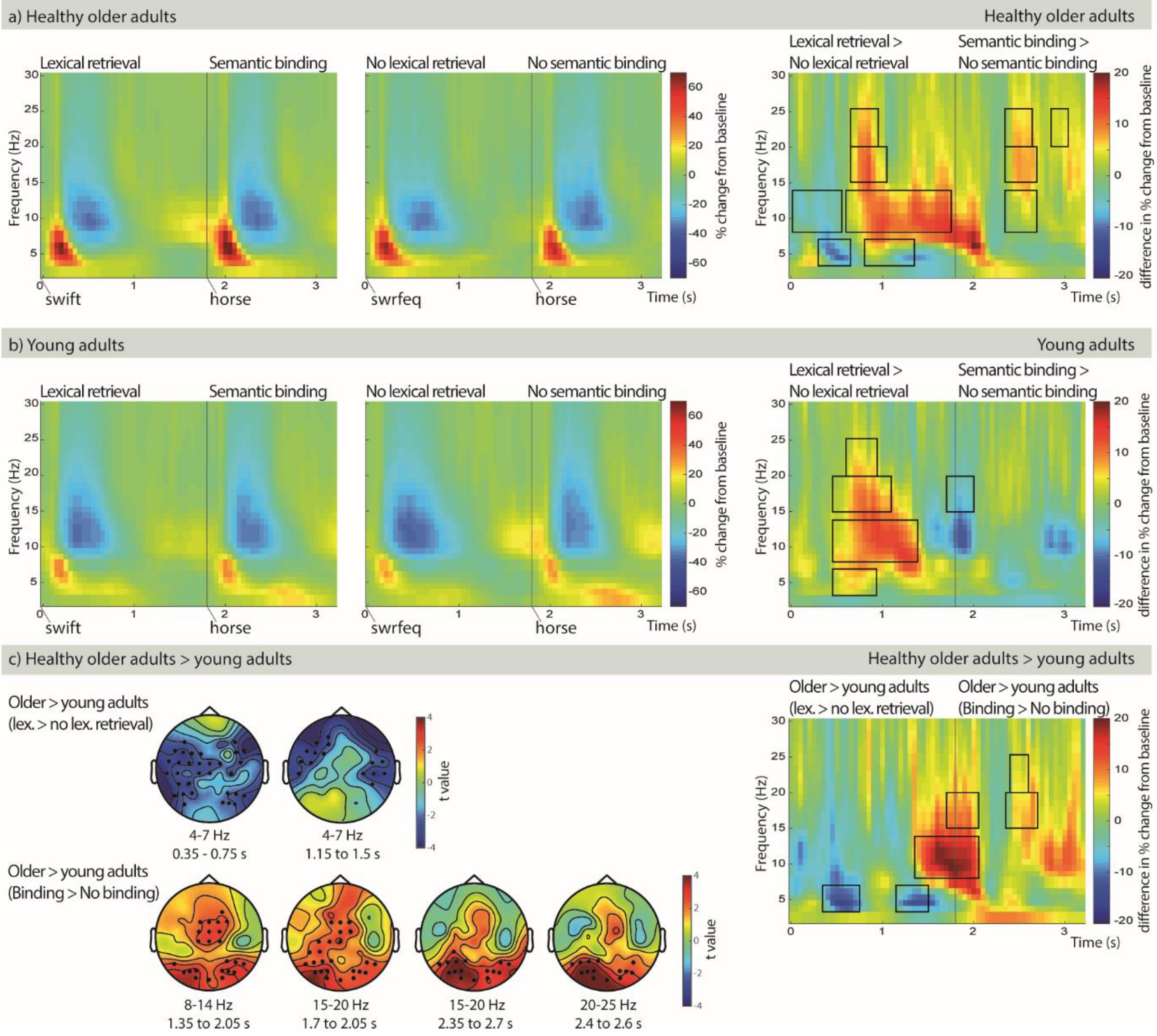


Figure 2.3 TFRs of power (collapsed across all electrodes) for lexical retrieval/semantic binding and no lexical retrieval/no binding, in (A) the healthy older adults, and (B) the young adults. (C) The condition differences (i.e. lexical retrieval/ semantic binding – no lexical retrieval/ no semantic binding) in older adults minus the condition differences in young adults. Head plots are illustrating the clusters of electrodes that show the most pronounced mean condition difference for the healthy older adults vs. the young adults. Black rectangles indicate significant group differences (*p*<0.05, cluster corrected).

2.5.1 Lexical retrieval results

We investigated the effect of lexical retrieval and recognition of the word form through comparison of power differences in the word 1 time window, comparing real words (i.e. lexical retrieval) to letter strings (i.e. no lexical retrieval can successfully take place (e.g., swift vs. swrfeq)). We observed significant differences between age groups for the lexical retrieval effects, and clusters in the observed data were found in the theta (4-7Hz) and alpha (8-14Hz) range (Figure 2.3C). We discuss these in turn below.

Theta and alpha power modulations for lexical processing show opposing patterns for healthy older adults and young adults

We observed significant differences in theta activity between the age groups (*p* = .014) during lexical retrieval (i.e. lexical retrieval minus no lexical retrieval) maximal over occipito-temporal electrodes. Specifically, we observed that theta was attenuated around 0.35 to 0.75 sec after the first word during lexical retrieval in the healthy older adult group relative to the young adult group. This age-group difference in theta activity during lexical retrieval emerged due to opposing patterns of theta modulation in relation to lexical and no lexical processing trials. Specifically, within the healthy older adults there was a smaller increase in theta power (*p* = .02) for lexical compared to no lexical retrieval in a corresponding time window, maximal over right occipital and left central electrodes (Figure 2.3A). However, in contrast the young adults showed a greater increase in theta power for lexical compared to no lexical retrieval (*p* = .02) in a similar time window (0.45 to 0.95 sec), maximal over left occipital and parietal electrodes (Figure 2.3B).

In addition, we observed differences in theta power related to lexical processing (*p* = .05) corresponding to a cluster extending between 1.15 to 1.50 sec, with the power being attenuated in the older adult group compared to the young participants. This age-group difference, pronounced over frontal-lateral electrodes, emerged due to an increase in theta activity in the non-lexical relative to the lexical condition (*p* = .004) in a cluster in the observed data that extended from 0.8 to 1.35sec in the elderly participants, which was absent in the young group.

Last, we observed a significant group difference (*p* = .034) in the alpha modulation (8-14Hz) that corresponded to a cluster extending from 1.35 to 2.05 sec. This observed cluster was maximal over occipital and central electrodes. This group difference emerged due to an opposite pattern of alpha power rebound during lexical processing between the older and young adults. In the older adults the alpha power rebound following the post-word alpha suppression was greater in the lexical retrieval trials than the no lexical retrieval trials (*p* = .002). A cluster in the observed data spanned from 0.6 to 1.75 sec. On the other hand a reversed pattern was seen in the young participant group with the alpha rebound being greater in the non-lexical retrieval trials compared to the lexical retrieval trials (note here that this is a qualitative description of the power modulations based on visual inspection only, rather than a significant effect). We interpret the attenuated alpha rebound to reflect an absence of closure in the no lexical retrieval condition in the older adults. Specifically, we hypothesize that the older adults continue to try and retrieve a lexical item (for a longer time than the young adults) after the onset of a pseudoword. Finally, it must be noted that the alpha rebound overlaps with the onset of the second word (i.e. the word to be semantically integrated). It is possible therefore that the between-group difference in alpha rebound must be partly attributed to semantic binding (discussed further below) rather than exclusively lexical processes.

Lexical retrieval results summary

Summarizing above significant between-group differences in lexical retrieval and recognition of the word form, we find that the older and young adults appear to exhibit opposite patterns of theta and alpha modulation after the onset of real words (i.e. lexical retrieval) vs letter strings (i.e. no successful lexical retrieval). Specifically, closely followed by word onset, older adults exhibited a smaller, whereas young adults exhibited a larger theta power increase, for the lexical compared to the non-lexical condition. Additionally, the alpha power rebound effect was reversed between the groups: in the older adults it was greater, and in the young adults smaller (though not significantly so), in the lexical retrieval condition compared to the non-lexical condition.

2.5.2 Semantic binding results

Semantic binding effects were defined by comparing oscillatory power surrounding the onset of a second (i.e. target) word in a semantic binding to a no semantic binding context (e.g., *horse* when preceded by *swift* vs. when preceded by *swrfeq*). We observed significant differences between age groups for the semantic binding effects, with clusters in the observed data being found in the alpha (8-14Hz), low beta (15-20Hz) and high beta range (20-25Hz) (Figure 2.3C). We discuss these in turn below.

Alpha and beta power in a time window preceding and following the onset of the target word was differentially modulated across the age groups

A cluster in the observed data was found in the alpha band (8-14Hz), which corresponded to a significant between-group difference (p = .034). The observed cluster extended from 1.35 to 2.05 sec over occipital and central electrodes. Although within-groups there were no significant condition differences, for young adults we observed that the alpha rebound was attenuated preceding and following the target word in the semantic binding condition compared to the no semantic binding condition. On the other hand for older adults in the equivalent time window an opposite pattern was observed with the alpha rebound being larger in the semantic binding relative to the no semantic binding trials. The pattern just described in the alpha range, extends into the lower beta range (15-20Hz). There was a significant between-group effect (p = .046) with a cluster extending from 1.7 to 2.05 sec, maximal over occipital, parietal and central electrodes. For young adults, there was a smaller low-beta increase in the semantic binding compared to the no semantic binding condition (*p* = .036), corresponding to a cluster that extended from 1.7 to 2sec.

A beta rebound was observed for semantic binding for the healthy older adults but not for the young adult group.

In addition, there were between-group effects that mapped onto clusters found in the lower beta (15-20Hz) and higher beta ranges (25-30Hz) over occipital electrodes that extended from 2.35 to 2.7 sec and 2.4 to 2.6 sec respectively. The between-group effect in the lower-beta (*p* = .002) and higher-beta range (*p* = .046) is driven by the older adults showing a clear semantic binding signature in this time window (with the effect extending into the alpha range), with no effect for the young adults in the equivalent time window. The healthy older adults elicited greater and more sustained suppression (*p* = .002) of lower-beta, and more higher-beta suppression (*p* = .002) in the no semantic binding condition compared to the semantic binding condition. These effects corresponded to clusters that extended from 2.35 sec and ended around 2.65 to 2.7 sec.

Semantic binding results summary

Summarizing above significant between-group differences in semantic binding, we see that young and healthy older adults clearly have different semantic binding signatures. In young adults, there is an attenuation of the alpha rebound in anticipation of the target word onset, followed by a binding signature in the low-beta band (i.e. a smaller low-beta increase in the semantic binding compared to the no semantic binding condition) immediately preceding and during the presentation of the target, to-be-integrated, word. In contrast, during the semantic binding condition the older adults exhibited a smaller decrease in high and low beta activity (compared to no semantic binding) starting only at 500 ms after the onset of the target word, which was not present in the young adults.

2.5.3 No effects of plausibility

We carried out a non-parametric cluster-based permutation analysis within the 1.8 to 3.2 sec time window of interest with pre-defined frequency bands (described above) to compare plausible semantic binding (e.g. swift horse) to implausible semantic binding (e.g. barking horse). We found no significant differences in the power changes in any of the pre-defined frequency bands between healthy older and young adults. Important is that we do not take this to suggest that there are no age-related differences between young and older adults in the effects of plausibility on semantic integration, but rather, that our experiment was potentially not suited to reveal them. We must note that within the older adults group, the condition comparison between plausible and implausible semantic binding elicited no power differences; within the young adults a power difference approached significance (p = .054), corresponding to a cluster that extended only from 2 to 2.15 sec. This observed cluster was found in the low beta (15 to 20Hz) band. It is likely therefore that our plausibility manipulation was not strong enough to elicit reliable condition differences, in either age-group, and therefore our experimental manipulation may not have been sensitive enough to investigate potential age-related changes in processing plausibility.

2.6 Discussion

The current study used a minimal two-word phrase paradigm to investigate the differences in oscillatory activity (in the theta, alpha, and beta range) during lexical retrieval and semantic binding in healthy older vs. young adults. Lexical retrieval was assessed by comparing neural patterns during the presentation of real words (e.g., swift) vs. letter strings (e.g., swrfeq). Semantic binding was examined by comparing neural patterns between semantic binding (e.g., horse, preceded by swift) and no semantic binding (e.g., horse, preceded by swrfeq) conditions. Here, it is important to highlight that although the present study manipulated the semantic binding context of the phrases, syntactic binding was also present in the semantic binding condition. Therefore, the semantic binding effects likely cover both semantic and syntactic compositional properties.

With regards to lexical processing we found that the older and younger groups exhibited opposite patterns of theta and alpha modulation at specific time intervals after word onset, which as a combined picture suggest that lexical retrieval is associated with different and delayed signatures in older compared to young adults. Interestingly, with respect to semantic binding, we observed a signature in the low-beta range for young adults (i.e. a smaller increase for semantic binding relative to no binding) surrounding the presentation of the target word, while the semantic binding signature for older adults occurred about ~500ms later as a smaller low- and high-beta decrease (for binding compared to no binding). We will now discuss each of these findings in more detail in relation to previous literature.

2.6.1 Age-related oscillatory patterns linked to lexical retrieval

Firstly, we found that the oscillatory patterns observed during lexical retrieval (i.e. post word one onset) were different for healthy older adults compared to young adults. The presentation of word one (regardless of condition) led to an increase in theta power across both age groups, which has previously been proposed to be linked with the role of long-term memory retrieval (Bastiaansen et al., 2002, 2008; Bastiaansen & Hagoort, 2006).

The lexical retrieval effect (i.e. a real word vs letter string) was associated with a greater theta increase closely followed after the first word onset (0.45 to 0.95 sec) in young adults over the left temporal and parietal sites, suggesting greater demand on retrieving the meaning of ‘real’ words from long term memory. In other words, the effort required to retrieve an item containing lexical information was greater compared to when the item lacked lexical representation. This is consistent with previous studies (Bastiaansen et al., 2005; Marinkovic et al., 2012; Mellem et al., 2013) who found that items that carry greater meaning or are more complex (i.e. real words and open class words) elicit stronger theta response compared to items that lack (or have lesser) lexical representation (i.e. pseudowords and closed class words) in young adults. In Shahin et al., (2009) participants made voice identification or semantic judgements to auditory word stimuli and found that theta power is heightened for the latter. Although the current study is visual, it appears that a similar mechanism is observed as for the auditory modality.

In contrast, amongst healthy older adults, the lexical retrieval effect was associated with a theta increase which was smaller closely after word one onset (0.3 to 0.65 sec) and then bigger (0.8 to 1.35 sec) in the lexical retrieval condition (vs. no lexical retrieval). In other words, the ‘typical’ lexical effect (i.e. greater theta increase in the lexical retrieval compared to no lexical retrieval condition) occurred ~350ms later in healthy older adults compared to young adults and was more widely spread. We suggest that healthy older adults require a longer time and a wider network to retrieve a lexical item from memory compared to young adults. However, it should be noted that this is a highly speculative conclusion as the topography of this lexical retrieval effect is varied between the age groups.

The lexical retrieval effect in healthy older adults in the theta band is contrary to Mellem et al., (2012) who did not find any theta power differences associated with lexico-semantic processing when comparing open and closed class words in older adults (although the lexical manipulations differ between the current study and Mellem et al. (2012), both paradigms manipulated the level of lexico-semantic content). Additionally, this theta power difference between lexical vs. no lexical retrieval was prominent over a more widely spread network in healthy older adults including the bilateral occipital and central electrodes compared to young adults where it was lateralised to the left temporal and parietal sites. Although our results do not permit us to definitively assert this resemblance, it echoes the frequently found tendency for healthy older adults to show a lesser engagement of task relevant regions but a greater involvement of other regions compared to young adults (Cabeza et al., 1997, 2002; Grady, 2000). Furthermore, the Hemispheric Asymmetry Reduction in older adults (HAROLD) model (Cabeza, 2002) suggested that the neural processing in healthy older adults is associated with a decrease in hemispheric asymmetry (whereas young adults show lateralization to one side), which is evident in the theta activity in the current study.

It should be noted here that linguistic processes other than lexico-semantics are involved in reading words vs. letter strings, including recognition of the word form and orthographic processing (Taylor et al., 2013); thus the results relating to the manipulation of lexical retrieval will also refer to these features. The left ventral occipito-temporal (vOT) cortex has previously been linked to word processing, specifically to aligning to orthographic features (Schurz et al., 2014). Froehlich et al. (2018) found that age-related differences in word processing were most pronounced during orthographic processing in (among others) the vOT circuit. This is consistent with our present age-related theta effect as this was most pronounced at the occipito-temporal sites. This suggests that the age-related difference in processing words vs. letter strings speculatively lies in the recognition of the word form and its orthographic properties.

Lastly, older adults did not show an alpha rebound like the young adults: while young adults showed an alpha rebound that was greater in the no-lexical retrieval compared to the lexical retrieval, the older adults showed the opposite. We interpret this to reflect an absence of closure in the no lexical retrieval condition in the older adults: older adults may continue to try and retrieve a lexical item (for a longer time than the young adults) after the onset of a pseudoword. In other words, processing of the non-lexical item was not fully complete, consistent with a finding in previous ageing studies using a paradigm with pseudo words (Poulisse et al., 2020). However, this interpretation is tentative and we elaborate on this in the next section.

2.6.2 Age related oscillatory patterns linked to semantic binding

Most interestingly, we observed oscillatory differences between healthy older and young adults during and after word two onset. The target word (i.e. the second word) in the semantic binding condition required participants to retrieve the lexico-semantic information from memory, and, to develop (i.e. bind together) a compound meaning representation of the two-word phrase. The latter was absent in the no semantic binding condition where the combination of a letter string and a real word cannot create a meaningful phrase. We observed clearly different semantic binding signatures (i.e. comparing the target words in the semantic binding vs. no semantic binding condition) for the young vs. healthy older adult groups. Although we refer to the below effects as semantic binding signatures, it is important to note here again that the comparisons of the respective conditions either simultaneously contained syntactic and semantic binding or no presence of binding at all. These are novel findings as few previous studies have investigated the effect of ageing on the oscillatory dynamics associated with semantic binding.

The semantic binding effect was associated with an oscillatory brain activity difference in the alpha (1.35 to 2.05 sec) and low-beta (1.7 to 2 sec) frequency bands between healthy older and young participants. The effects that occur around the onset of the second word must be treated with caution. Due to the caveats of the experimental design, one cannot easily map these effects onto either the lexical status of the first word or the anticipatory processing related to binding of the second word. We can only present plausible explanations for the observed results based on the comparisons to previous literature. Therefore, here we speculatively propose that the first part of the group difference in the alpha band (up until around 1.7 sec) occurred due to older adults eliciting an atypical alpha rebound response in the no lexical retrieval condition, whereas the young adults had no condition effect in the same time window (already discussed in the previous section on lexical retrieval). The later part of this alpha effect (i.e., post second word onset) may potentially be implicated in the binding process itself, including making semantic predictions. Although previous research (e.g., Luo et al., 2010) has related alpha power modulations with a violation of semantics (and therefore binding), these effects occur 400-600 ms after the onset of the target word and not during it. For this reason, we cannot definitively conclude whether this alpha power modulation surrounding the onset of the second word is related to the lexical retrieval closure of the first word, the anticipatory processing of binding, or a combination of the two.

Importantly, young adults elicited a smaller low-beta increase in the semantic binding condition (vs. no semantic binding) in a time-window immediately preceding and during the presentation of the to-be-integrated target word (from 1.7 to 2 sec for low-beta). This early binding signature in the young adults within the beta band is somewhat consistent with previous studies. Beta frequencies have previously been proposed to be “carriers” of linguistic information and are involved in binding past and present inputs (Weiss & Mueller, 2012). For example, von Stein et al., (1999) found coherence exclusively in the beta frequency range between left temporal and parietal sites during semantic binding across visual and auditory modalities. Additionally, Berghoff et al., (2005) showed that figurative compared to literal sentences elicited increased coherence in the beta band between the hemispheres during the binding of semantically related information.

On the other hand, our findings are contrary to the linguistic information maintenance theory (Lewis & Bastiaansen, 2015). They proposed that any changes in the processing of the linguistic input (i.e. violations of semantics or syntax) lead to greater beta power desynchronization. However, it is important to note here that in our experimental design we do not induce semantic binding violations but rather compare semantic binding vs. no semantic binding. This may partially explain why the current effects differ from some previous empirical studies. For example, Luo et al., (2010) showed that semantically incongruous sentences elicited a reduced beta power 0 to 200 ms post presentation of the critical word (compared to congruous sentences). The time window of the beta modulation of our results (i.e., immediately after the target word onset) is similar to Luo et al. (2010) indicating that this effect may be related to the anticipatory and prediction processes involving binding of the second word. Furthermore, the direction of the beta power modulation is opposing to our findings as during this time window the young adults displayed a reduced beta power increase in the semantic binding condition (equivalent to the semantically congruous sentences) compared to the no binding condition (similar but not equivalent to the semantically incongruous sentences). Additionally, Lewis et al. (2017), also found a greater beta desynchronization in the incoherent condition (compared to coherent condition) when investigating the effect of semantic coherence at a local sentence level, when using short stories. However, it is important to highlight here that this semantic coherence manipulation was only effective at significantly modulating the beta power at the last sentence presentation (in a story with 4 sentences). Lastly, Wang, Jensen et al. (2012) found that sentences ending in anomalous words induced a beta power decrease (in the same time window as the N400) over the left temporal areas. These beta modulations significantly correlated with the N400: a larger beta power decrease was associated with smaller N400 amplitudes. The authors suggested that the role of beta oscillations within language comprehension is complex, but it is evident that beta oscillations are involved in semantic unification of items into the wider phrase or sentence context. At this point, the reasons for our somewhat conflicting findings remain unclear. Although we do not know exactly how the current results relate to previous findings, it is crucial to communicate them as this will allow for theories to be updated and developed further.

Compared to the young adults, the binding signature in the healthy older adult group was different, and moreover delayed by about 500 ms. Healthy older adults elicited a smaller beta decrease in the semantic binding condition (vs. no semantic binding). This semantic binding effect in healthy older adults is consistent with Meltzer et al. (2017) who observed a greater magnitude of 8-30Hz event-related desynchronization for word lists (no semantic binding condition) compared to sentences (semantic binding condition) in older adults. However, they also observed this effect in young adults, and therefore it may be surprising that we do not see this condition difference in the young group in an equivalent time window. Furthermore, the binding effect in healthy older adults is in line with the maintenance theory by Lewis and Bastiaansen (2015) and its supporting empirical evidence. The no semantic binding condition led to a greater low beta power decrease ~500 ms after the onset of the target word compared to the semantic binding condition as the language comprehension system detected the need for change. This binding effect in the healthy older adults follows a similar timing pattern as Luo et al. (2010), whereas the young adults did not show any significant differences between the two conditions (binding vs. no binding), implying that surprisingly the language comprehension system did not detect any requirements for change in processing. If the early beta power modulation (during the onset of the second word) in young adults is related to the anticipatory/ prediction activation, the system does not need to ‘listen out’ for any changes later on as it has already predicted the linguistic outcome. It is possible that healthy older adults are sensitive to the requirement for the system to change its processing (more so than young adults) as they were not able to anticipate or predict the incoming binding during the onset of the second word. Though, this is only speculation. This late binding effect in the beta band in healthy older adults is not fully in line with the maintenance hypothesis (Lewis & Bastiaansen, 2015), as we see a beta decrease in both semantic binding and no binding conditions (just a smaller one than in the no semantic binding).

The binding signature we see for healthy older adults is not only delayed but also in the opposite direction to that of the young adults. The finding of an inverse effect between a young and older adult age group is similar to some previous studies. For example, Beese et al., (2019) in an auditory sentence comprehension study tested whether oscillatory power differed across age groups when comparing correctly and incorrectly encoded sentences. They reported that young adults displayed a negative effect (later remembered vs later not remembered sentences produced an alpha decrease). This effect was attenuated in the middle aged adults and completely inversed in older adults (later remembered vs later not remembered sentences produced an alpha increase). The authors attributed this alpha band effect to a shift from cortical disinhibition to inhibition during sentence encoding. Additionally, Poulisse et al., (2020) also found an inverse condition effect between older and young adults in a syntactic binding context (in a two-word phrase paradigm). However, their inverse effect was contrary to our findings (i.e. they found that the syntactic binding effect was associated with a larger alpha power increase in young adults, and a smaller alpha power increase amongst the older adults). Although this pattern is the opposite of our results, our paradigm manipulated semantic and not syntactic binding. Also, it is important to note that Poulisse et al., (2020) observed this inverse condition effect between the groups in a much later time window (0.6-1.05sec) after the phrase presentation compared to our study (0-0.3sec).

As changes in the beta frequencies echo the role of language-related binding (Weiss & Mueller, 2012) we summarize that the semantic binding signature (reflected in the beta frequency) occurred *during* the presentation of the target word (which is when the semantic binding process takes place) in the young adults and was delayed by ~500ms in the healthy older group. However, these conclusions need to be treated with caution, as the interpretations of our results are speculative.

2.6.3 Suggestions for future research

We found different neural signatures in oscillatory power for young vs. healthy older adults in the present study. Differing functional neural patterns in healthy older adults are commonly interpreted as being compensatory (e.g., Cabeza et al., 2002). However, the term compensation should be reserved for differing neural patterns which are contributing meaningfully to performance (Cabeza et al., 2018; Grady, 2012). The present study was not designed to relate changes in brain function to language performance. Ideally, future studies would incorporate a trial-by-trial semantic comprehension performance measure and characterize which changes in network dynamics are predictive of successful language performance. Investigating the direct relationship between age-related functional neural changes and behavioural performance would be a particularly interesting avenue for future research, and be a necessary step in answering the fundamental question of how the ageing brain adapts to structural decline and reorganises its mechanistic functioning to support language comprehension.

In the present study, we used an experimental paradigm that focused on semantic binding, while minimizing contributions of the ability to make predictions. However, natural language comprehension does rely to some extent on actively making predictions. Making predictions gives language processing a head start (Kuperberg & Jaeger, 2016). Previous studies on young adults have found an alpha power decrease prior to the onset of predictable words (Rommers et al., 2017; Wang et al., 2018), suggesting that in young adults anticipatory binding processes are initiated prior to predictable words being presented. A number of ERP studies have demonstrated that there are age-related changes in the ability to engage prediction mechanisms. Age-related changes in processing predictable sentence endings were evident by the lack of a frontal positivity effect for older (compared to young) adults (Wlotko et al., 2012). Also, N400 amplitudes suggest that older adults do not use the sentence context to pre-activate semantic features of predictable words (Federmeier et al., 2002). Future studies could extend on these findings and examine the age-related changes in oscillatory dynamics supporting the use of predictions during semantic binding.

Lastly, future research should incorporate individual differences measures (and thus, larger sample sizes) to assess which non-linguistic cognitive resources and brain structure properties support the implementation of age-related changes in functional neural characteristics (Peelle, 2019). A quantitative shift in capacity constraints (e.g. higher working memory) can qualitatively impact on the way language is processed, for example, making older adults more able to use contextual semantic information or predictions. Moreover, sufficiently flexible cognitive resources can work together to circumvent structural decline and support functional adaptations, maintaining successful language and communication performance.

2.6.4 Summary

Healthy older adults have a different oscillatory signature for semantic binding compared to young adults: young adults elicit an early semantic binding signature, around the target word presentation, in the form of a smaller low-beta increase during semantic binding (compared to no semantic binding). On the other hand, healthy older adults display a semantic binding signature ~500 ms later, with a smaller low/high-beta decrease in the semantic binding condition (compared to no semantic binding). Our findings are in line with previous literature suggesting that older adults do extract and make use of contextual semantic information, but there are differences (compared to young adults) with respect to when and how this happens (Federmeier et al., 2003; Federmeier & Kutas, 2005; Stine-Morrow et al., 1999; Wlotko & Federmeier, 2012).

CHAPTER 3: BRAIN-TO-BRAIN COUPLING FORECASTS FUTURE JOINT ACTION

In this study, we investigated whether brain-to-brain coupling patterns in a two-player time-estimation game could predict performance in a cooperation task. The participant pairs (N=30), were tasked with synchronising button presses after converging (and waiting) on a shared representation of 'short', 'medium', and 'long' time intervals while utilizing feedback to adjust responses. We employed EEG-hyperscanning and focused on post-feedback brain activity. We found that negative feedback led to increased frontal mid-line theta activity across individuals. Moreover, a correlation in post-feedback theta power between players forecasted failed joint action, while anti-correlation forecasted success. These findings suggest that temporally coupled feedback-related brain activity between two individuals serves as an indicator of redundancy in adjustment of common goal representation. Additionally, the anti-correlation of this activity reflects cognitive strategic mechanisms that ensure optimal joint action outcomes. Rather than a *paired overcompensation*, successful cooperation requires *flexible strategic agility* from both partners.

Published in:

Markiewicz, R., Segaert, K., & Mazaheri, A. (submitted). Brain-to-brain coupling forecasts future joint action. *Pre-print available on:* https://doi.org/10.1101/2023.09.21.558775

Data availability:

Stimuli and data are available here: https://osf.io/ct8jb/

3.1 Introduction

Successful joint action entails sharing a common goal and striving towards it, while also incorporating feedback when deviations from the goal occur. To successfully cooperate, it is thus crucial to process feedback and adjust behaviour accordingly. While EEG-hyperscanning has previously been used to investigate joint action (Hamilton, 2021), no study to date has investigated the brain-to-brain coupling of neurocognitive mechanisms that drive the adjustment of shared representations of the common goal. In the current study, we investigate the specific brain-to-brain activity, in response to negative feedback, that can forecast the strategic cognitive mechanisms underlying adjustment of the shared representation of the task and facilitate successful performance in joint action cooperation task.

The primate's brain has evolved to not only manage the demands related to cognitive abilities but more importantly to navigate and work with large social groups (Dunbar, 1992). Joint action is a type of social interaction, in which a shared goal is achieved through the coordinated actions of at least two individuals in time and space (Sebanz et al., 2006). Joint action is present in everyday life, for example when moving a heavy sofa up a flight of stairs. The success of joint action performance is reliant on creating shared representations, the ability to predict actions as well as the outcomes of one's own and others' actions (Sebanz et al., 2006). A first mechanism that underlies joint action is about ensuring that the coordinating partners can *guide attention* to perceive the same event or object (like creating common ground). Secondly, for successful joint action to occur, *action observation* is required. A corresponding representation of the object/event of interest in the observer's action system is created. This aids the understanding of the action, creates 'common ground' in terms of action goals, and supports prediction of each other's action outcomes. For example, Flanagan and Johansson (2003) observed that the gaze of the action observer precedes the action of their partner. Lastly, but perhaps the most important, is *action adjustment,* which refers to adjusting behaviour in space and time to complement the action of the coordinating partner. For example, lowering your end of the sofa when you see your moving partner is getting stuck with their end (i.e., the results of the adjusted shared representation of the common goal). Joint action is influenced not just by an individual's beliefs about their own abilities but also by the beliefs about what they can achieve in collaboration with others (Marsh et al., 2006). In the present study, we will focus on using feedback to make the adjustments needed to increase joint action success. To do this, we will examine the dynamic relationship between interacting brains (Dumas, 2011; Hari et al., 2015; Konvalinka & Roepstorff, 2012; Stolk et al., 2016) using a hyperscanning approach.

In the last decade, the social neuroscience field has seen a trend towards using a 'multi brain framework' to study how humans interact with each other. Hyperscanning involves the simultaneous recording of brain activity (either hemodynamic or neuroelectric) in multiple individuals during interactive task performance. It allows for investigations on how the activity dynamics of two brains (or more) underlies the continuous adaptation of one's actions in response to the changes in actions of someone else. Since first implemented by Montague et al. (2002), two-brain science has been used to study classroom dynamics (Dikker et al., 2017), communication (Stolk et al., 2013), music (Babiloni et al., 2012), and coordinated button presses (Balconi & Vanutelli, 2017b; Cui et al., 2012; Funane et al., 2011) (see Czeszumski et al., 2020; Konvalinka & Roepstorff, 2012 for detailed reviews). Also interpersonal dynamics such as cooperation or joint action have widely been studied using hyperscanning methods (see Balconi & Vanutelli, 2017a for detailed review). Joint action relies on the postulation of anticipating and predicting other people's actions and adjusting our own actions accordingly (Decety & Sommerville, 2003). 'Coupling' between individuals can occur on motor, perceptual or cognitive levels during joint action (Knoblich et al., 2011) with behavioural synchrony often reflected in neural synchrony. This neural connectivity as measured by hyperscanning is the correlation between neurophysiological signals that are temporally related but spatially distinct (i.e., in two separate brains) (Dumas et al., 2012).

EEG hyperscanning is particularly useful in the context of the present study due to its high temporal resolution, which allows for a real-time window into the neural processes induced by feedback in the unfolding dynamics of social interaction on a millisecond scale. Several EEG hyperscanning studies have reported inter-brain synchronous oscillatory patterns in joint action contexts, both in terms of oscillatory power as well as phase during fine motor coordination (Dumas et al., 2010; Tognoli et al., 2007), verbal coordination (Ahn et al., 2018) and other complex joint actions (Astolfi et al., 2012). The mechanisms that support adjustment of the shared representation of the task goal in response to feedback in a joint-action action context are key to explaining joint action outcomes, but they are not clear to us yet. Only two studies to our knowledge have investigated the inter-brain dynamics in tasks with an incorporated feedback element. Mu et al. (2016) reported on the heightened inter-brain alpha phase synchrony *during* the coordination task (vs control), but not in response to feedback. Conversely, Balconi et al. (2018) examined directly the effect of external feedback on joint action in the context of EEG inter-brain dynamics. However, in this study feedback was superficially created and not based on the actual joint action performance. From this, it is evident that there is a clear gap in the literature: the precise neural mechanisms that facilitate adaptive responses to feedback and thus shared representation of the common goal within the context of joint action remain elusive.

The impact of feedback has been studied extensively in tasks that do not involve a joint-action component (e.g., van de Vijver et al., 2011), or, in tasks that involved a joint-action component but only examined the neural dynamics of one of the participants (e.g., Czeszumski et al., 2019; Itagaki & Katayama, 2008; Picton et al., 2012). A careful look at these studies enables us to propose precise predictions about the inter-brain mechanisms that might be involved in adjusting based on feedback in a joint-action setting. It has been established that a feedback cue induces oscillatory EEG changes, time-locked, but necessarily phase-locked, in the 3-7 Hz (theta range) with a maximal midline frontal distribution. This feedback induced neural signature is thought to reflect the initiation of response adjustments (i.e. cognitive control) overriding ‘status quo’ responses (Cavanagh et al., 2013; Cavanagh & Frank, 2014). The midline theta response can be reliably measured on a trial-by-trial basis, providing valuable insights into the neural-dynamics subserving behavioural adjustment. This measure affords a real-time window into how the brain adapts and adjusts behaviour based on feedback (van de Vijver et al., 2011), allowing us to explore the intricate relationship between neural activity and behavioural responses. In the present investigation, our focus lies on theta activity as a reflection of adjustment in the mental representation of the time intervals between the partners within a pair. However, rather than solely examining how the amplitude of feedback-related theta activity in each participant (referred to as "player") relates to their performance in future trials, we, by using EEG hyperscanning, are specifically interested in how the *correlation* between these responses can effectively forecast their joint action success.

We implemented an EEG hyperscanning approach while participants completed an innovative time-estimation task. The objective of this task was to converge on the mutual representation of what a ‘short’, ‘medium’ and ‘long’ interval entailed. Specifically, auditory cues instructed participants to either press a button after what they mutually judged to be a short, medium or long time. Participants reached a successful joint action outcome if they reached a shared representation of the time intervals and thus converged on their responses. Importantly, feedback was given after each trial with respect to their joint action outcome, and if unsuccessful, about who was faster and by how much. With this, we created an ideal situation to study the inter-brain mechanisms that support feedback processing and subsequent behavioural adjustment. By utilizing this time-estimation task, we were able to study how individuals develop shared understanding and mutually converge on an idea (in this case of timing intervals) exclusively through the external feedback they received. This task included well-defined joint action outcomes, namely successful and failed cooperation, enabling us to investigate the separate neural mechanisms underlying these distinct outcomes in joint action. Specifically, on an individual level, we seek to explore the degree to which the feedback-related theta signature extends to a joint action task. Secondly, and of particular interest, we aim to investigate the inter-player theta power dynamics in response to negative feedback, with a focus on anticipating cooperative outcomes in a joint action task. This will involve a comparison of brain-to-brain theta power dynamics in trials *categorized* as either forecasting successful or failed cooperation. Additionally, we will explore the cooperative outcome as a continuous variable, examining whether the interaction between the theta power of player one and player two can predict the cooperative outcome, measured as the raw reaction time difference between their button presses.

In line with previous research (e.g., van de Vijver et al., 2011), at the individual level, we expect a greater increase in frontal-midline theta power in response to negative feedback compared to positive feedback. Further, we anticipate that the interplay of feedback-related theta power of player one and player two would serve as a predictor of the cooperative outcome. Specifically, in our theoretical framework, we hypothesize that the absence of strategy alteration or mutual adjustment in the mental representation of the time intervals (short, medium, and long) in both players, referred to as *strategic stagnation*, would result in unsuccessful joint action. Conversely, if both partners actively modify their strategies (i.e., mutual increase in theta power in both players and thus correlated brain-to-brain activity), adapt their mental representations of the time intervals, this could potentially facilitate adaptive adjustments and promote successful joint action.

However, the risk of both players attempting to adjust their strategies is a redundancy of adjustment in the mental representations of the time intervals, which could lead to *paired overcompensation* and thus maladaptive behavioural adjustment. Therefore, our alternative hypothesis posits that joint action success is more likely when players exhibit *flexible strategic agility*. In this scenario, one partner adjusts their mental representation to a greater extent than the other, allowing for asymmetry in their adaptation process (i.e., anti-correlated brain-to-brain activity in the induced theta power). This asymmetry may manifest as one player making more pronounced changes in their mental representation compared to the other, leading to effective coordination and ultimately contributing to successful joint action. We hypothesise that this mechanism may strike an optimal balance, characterized by flexible strategic agility and asymmetrical adjustments in the mental representation of time intervals, and thus enable a balanced mutual convergence that facilitates successful joint action.

3.2 Materials and Methods

3.2.1 Participants

Sixty young adults (i.e., 30 participant pairs) took part in the study. These were University of Birmingham (aged 19 to 31, *M*: 21.5, *SD*: 2.3; 12 males) students who were compensated for their time with cash payments. One participant was excluded from all analyses due to excessive EEG artefacts in the recordings; while their cooperative partner's data were excluded from the *paired* analysis, they were kept in for the *individual* analysis. Subsequently, we included 59 participants in the *individual* analyses, and 29 pairs for the *paired* analyses.

All participants were native English, monolingual speakers, right-handed, with normal-to-corrected vision, and no neurological or language impairments. All participant pairs reported to not know their cooperative partner in the experiment. Participants signed informed consent, which followed the guidelines of the British Psychology Society code of ethics and the study was approved by the Science, Technology, Engineering, and Mathematics (STEM) Ethical Review Committee for the University of Birmingham (Ethics Approval Number: ERN_19_1661).

3.2.2 Cooperation task

Cooperation was measured via a non-verbal task, completed in pairs (see Figure 3.1A for experimental set up). Participants were instructed to synchronise button presses with their partner within the pair. One of three types of auditory stimuli (see Figure 3.1B) was played simultaneously to both participants at the start of each trial using headphones (Sennheiser 289 HD). These stimuli indicated the duration of time (referred to as ‘short’, ‘medium’, or ‘long’) that participants had to wait before pressing their button and attempting to synchronise it with the other participant in the pair. Neither of the participants within a pair knew the representation of ‘short’, ‘medium’, or ‘long’ the other participant had. There was no verbal or non-verbal communication between participants to agree on a mutual representation of these time intervals - they adjusted their responses solely based on the feedback they received at the end of each trial.

Cooperation was considered successful when both participants pressed their buttons within a 250ms timeframe. On the other hand, a trial was classified as failed cooperation if the

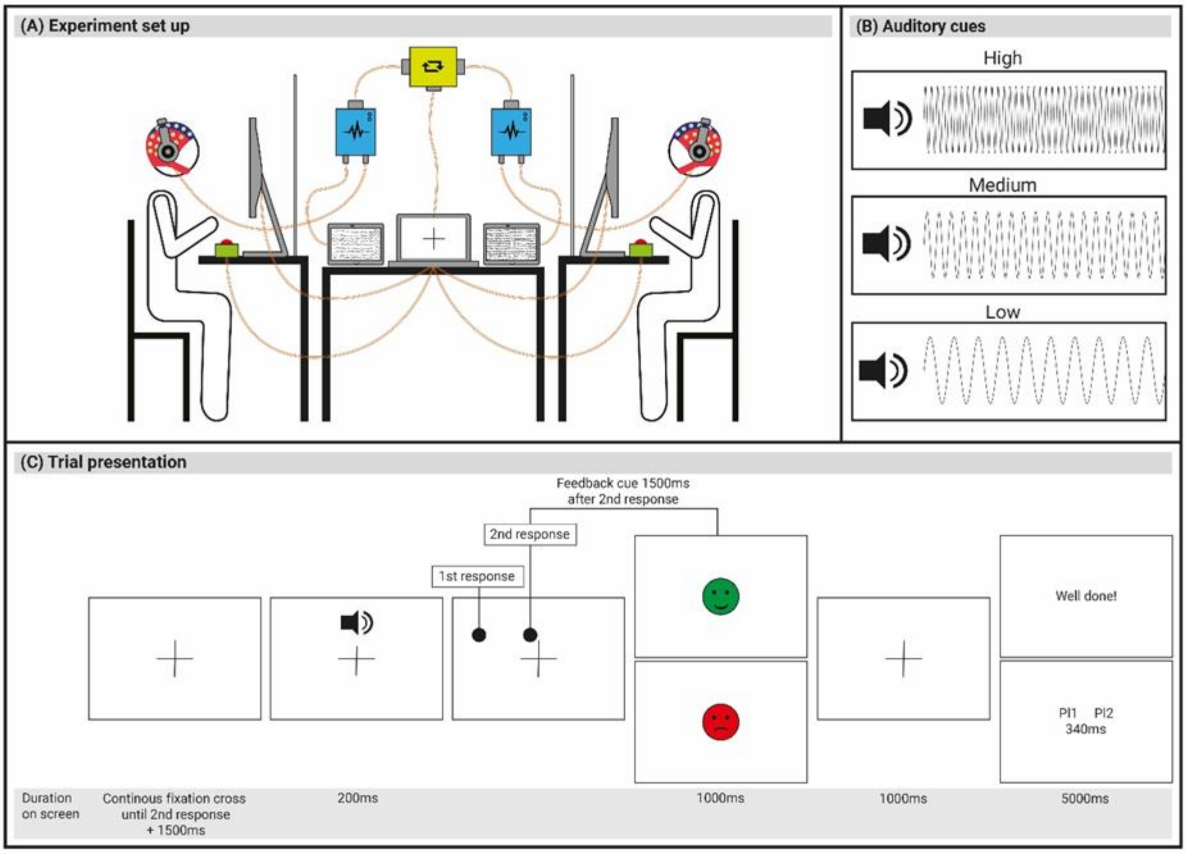


Figure 3.1 The experimental setup and trial presentation. (A) Participants wore Waveguard caps with 32 cap-mounted Ag/AgCI electrodes. The ANTneuro EEGsports amplifier systems amplified the signals for each participant. The EEG signals were synchronized using a parallel port splitter. The stimuli were presented on a central laptop connected to two identical monitors. Participants listened to auditory cues through noise-cancelling headphones and used a button box to record their responses. (B) The experiment used three auditory cues: high, medium, and low frequency tones, corresponding to short, medium, and long waiting times before pressing a button after what they mutually judged to be a short, medium, or long time. (C) Trial presentation involved both participants receiving the same stimuli simultaneously. The order of the button responses was completely driven by the participants and thus resulting in trial-to-trial variations of who pressed first/ second. The feedback stimuli varied based on trial outcome: a green smiley face and "Well done" for successful cooperation trials, and a red sad face with feedback text indicating which player pressed the button first/second and the elapsed time for failed cooperation trials. Participants were informed of their player number to understand the feedback text.

time elapsed between the two button presses was <250ms. Participants were provided with feedback regarding their performance, as explained below. We selected this timeframe threshold after conducting piloting of the paradigm (using a different set of participants) to attain an accuracy rate of approximately 50% across conditions. The purpose of setting the task goal with this threshold was to maintain an approximately equal number of trials per condition, including successful and failed joint action. An analysis of the accuracy rates in the present dataset revealed that we succeeded in our purpose: the overall accuracy across all pairs was 56.39% (SD = 13.94). Moreover, the accuracy rates revealed that the task worked as expected, with accuracy being lower to more difficult the condition was: A Repeated measures ANOVA (Greenhouse Geisser adjusted) revealed that there was a significant effect of condition type (high/short, medium, and low/long) on accuracy (%) ($F(1.66, 46.469) = 76.638$, $p < .001$). Post-hoc Least Significant Difference (LSD) pairwise comparisons showed that the accuracy rates were significantly different for each condition (high vs. low, $p <.001$; high vs. medium, $p < .001$; medium vs. low, $p = .037$), with the high/ short condition having the highest accuracy (M = 74.586 , SD = 14.151), followed by the medium condition (M = 49.379, SD = 18.108), and with the low/ long condition having the smallest accuracy (M = 45.207, SD = 15.697).

The experiment was set up using Python 3.6 and incorporated in-house built scripts and PsychoPy functions. The experiment scripts can be downloaded from the following link: https://osf.io/ct8jb/. Both participants viewed identical instructions, stimuli, and trial presentations on separate screens with a resolution of 1920 x 1080. The button presses were recorded using Razer gaming keypad. This gaming pad uses high-precision mechanical keyboard switches and specialized internal circuitry that purportedly enables the keyboard state to be polled at a rate of 1000 times per second.

Each trial commenced with the appearance of a fixation cross, which remained on the screen for the entire duration of the trial. After 300ms, an auditory cue was played for 200ms to both participants. There were three distinct auditory cues used: (1) a high-frequency 'beep' (2500 Hz), (2) a medium-frequency 'beep' (1000 Hz), and (3) a low-frequency 'beep' (200 Hz). Participants were instructed to wait for a 'short' amount of time upon hearing the high tone, a 'medium' amount of time upon hearing the medium tone, and a 'long' amount of time upon hearing the low tone. See Figure 3.1C for trial presentation.

Following the auditory cue, participants were required to press their buttons and synchronize their actions. Once both participants had made their button presses, the fixation cross remained on the screen for an additional 1500ms, after which a feedback symbol was displayed. Each trial had two potential outcomes: (1) successful cooperation/joint action, indicated when the elapsed time between button presses of participants within a pair was ≤250ms, or (2) failed cooperation/joint action, indicated when the elapsed time between button presses of participants within a pair was >250ms.

For successful cooperation, both participants were shown a green smiley face as visual feedback, while failed cooperation was represented by a red sad face for both participants. This visual feedback remained on the screen for 1000ms, followed by the presentation of a fixation cross for 1000ms. Subsequently, a feedback text was displayed to both participants for 5000ms. The text varied depending on the outcome of the trial, with participants receiving either a 'Well done!' message if their button presses were synchronized, or feedback indicating which participant pressed the button first/second and the time between the button presses (see Figure 3.1C).

An example of a trial presentation is depicted in Figure 3.1C. The experimental task consisted of 300 trials in total, divided into 10 blocks (30 trials in each block). Each block consisted of an equal number of the different auditory cues which were presented in a random order (10 trials with a high beep, 10 trials with a medium beep, and 10 trials with a low beep). In total, the experiment included 100 trials of each auditory cue. Participants were given the opportunity to have a break after each block.

3.2.3 Procedure

Each experimental session started with the 32-electrode EEG set up of each participant. Once both participants were capped, they sat approx. 70cm from a monitor (opposite each other with a table divider in between them so that they could not see each other) (see Figure 3.1A for set up). Participants wore earplugs (this was to prevent hearing each other's button presses) as well as noise cancelling headphones (to hear the auditory cues). Participants read the instructions of the task and once both participants were happy to continue, the experimenter started the 'training' phase.

The training phase involved participants hearing each of the auditory tones (high, medium and low) to familiarise themselves with the tones. While each tone was played, the associated description (i.e. high, medium, low) of the tone was displayed on the screen. This was repeated twice to ensure that participants were able to distinguish between the tones.

The experimenter was not present in the room during the experiment, instead the experimenter observed the participants through a one-way mirror. The experimenter entered the room upon the completion of each block and ensured that both participants were happy to continue. The experimenter controlled the start of the next block after each break.

3.3 EEG recording

The EEG set up was the same for both participants within a pair. EEG was recorded using Waveguard caps containing 32 Ag/AgCI electrodes (10-20 layout). The EEG signal was acquired with online reference to the CPz channel. The ANTneuro EEGosports amplifier system was used to amplify the signal and the EEGo sports was used to record it. The recorded signal was sampled at 500Hz, with a 150 Hz low pass filter and a 0.05Hz high-pass filter. We aimed to keep the impedances below 10kΩ. The EEG signal of both participants was synchronized to the onset of the triggers using a parallel port splitter.

3.4 EEG analysis

3.4.1 EEG pre-processing

The EEG preprocessing was performed using EEGLAB (2019.0) (Delorme & Makeig, 2004) and Fieldtrip toolbox (2021-10-16) (Oostenveld et al., 2011). The preprocessing steps for both participants in each pair were identical. Firstly, the first 5000 data points were removed from the continuous EEG signal to remove the amplifier start up noise. The continuous signal was then filtered using a high pass filter of 0.1Hz. The notch filter was then used to eliminate the electrical noise as well as its harmonics (45 to 55Hz; 95 to 105Hz; and 145 to 155Hz). The continuous data were visually inspected, and any noisy channels were removed and then interpolated (M = .27, SD = .67, max = 2). The signal was then re-referenced to the average of all electrodes and the CPz electrode (the online reference channel) was then added back into the data.

The data were then epoched to the onset of the feedback face (-7 to 7sec) (fourth screen in Figure 3.1C). Ocular artefacts were removed based on the scalp distribution using independent component analysis (ICA) in EEGLAB (2019.0). To improve the signal-to-noise ratio and classification accuracy of components, the EEG signal was high pass filtered at .8Hz (Winkler et al., 2015) for nine recordings prior to conducting ICA. ICA weights obtained from the filtered data were then added back to the original signal (filtered at 0.1 Hz). The average number of removed components was 2.88 (SD = 1.31) for each participant. For the subsequent analysis, the data were further epoched to the onset of the feedback text (the last screen of Figure 3.1C) (-4 to 4.8sec) and trials were sorted into meaningful conditions depending on the type of analysis carried out (see below).

3.4.2 Feedback-related analysis

To compare neural activity related to feedback after successful and unsuccessful cooperation, we conducted Time-Frequency representations (TFRs) of power and Event-Related Potentials (ERPs) analyses across all 59 participants. The successful condition consisted of trials (mean N = 155.64, SD = 41.08) in which both players within a pair pressed their buttons within 250ms of each other. The unsuccessful cooperation condition consisted of trials (M = 120.89, SD = 40.21) in which the time elapsed between the button presses of the players was greater than 250ms.

3.4.2.1 Time-Frequency representations of power analysis

The TFRs of power were calculated using the ‘mtmconvol’ method in Fieldtrip. In line with previous work (e.g., Markiewicz et al., 2021) we used sliding Hanning tapers with an adaptive time window of 3 cycles per each frequency of interest. The frequency of interest range was 2-50Hz in steps of 1Hz, and the time of interest of -1 to 4.8sec in steps of .05sec locked to the onset of the feedback text of each trial. The TFRs were calculated in relation to the relative change in power from baseline, with the baseline period being -0.6sec to -0.1sec (locked to the onset of the feedback text).

The statistical significance of the differences between conditions (successful vs. unsuccessful cooperation) in feedback cue locked oscillatory power changes, specifically locked to feedback cues in the time-frequency domain, was evaluated using non-parametric cluster-based permutation tests (Maris & Oostenveld, 2007). Each channel/time/frequency data point pair (successful and unsuccessful cooperation) locked to the onset of the feedback text was compared using a dependent samples T-test with a threshold of a 5% significance level. In cases where these comparisons (successful vs. unsuccessful cooperation) exceeded the significance level, the significant pairs were then clustered based on spatial proximity (via the triangulation method – a minimum of two neighbouring electrodes were considered a cluster). The Monte Carlo simulation was then performed to obtain probability values for the clusters by randomly shuffling the condition labels 1000 times and calculating the maximum cluster level test statistic for each permutation. The permutation tests were carried out on the following pre-defined and averaged frequency range bands: theta (4-7Hz), alpha (8-14Hz), low beta (15-20Hz), and high beta (20-25Hz), using a time window of 0 to 4 seconds locked to the onset of the feedback text.

3.4.2.2 Event Related Potentials analysis

The ERPs were calculated by averaging the time-locked EEG activity of all trials, separately for successful and unsuccessful cooperation conditions for each participant. The baseline correction used for the ERP analysis was -0.1sec to 0 prior to the onset of the feedback text. We employed the non-parametric cluster-based permutation tests (see above for full explanation; for comparison of ERPs the frequency part is not applicable) to statistically examine condition differences (successful vs. unsuccessful) in ERP amplitudes in a 0 to 4sec time window, locked to the onset of the feedback text. See supplementary material and Suppl. Fig. SM.3 for ERP results.

3.5.1 Forecasting cooperative outcome analysis

3.5.1.1 Effects of brain-to-brain coupling on local cooperative outcomes analysis

To identify power brain-to-brain signatures between participant pairs that were predictive of *local* cooperative outcomes, we focused our analysis of the TFRs of power locked to feedback in failed joint action trials. Importantly, the TFRs analysis was performed on the same set of trials for both player one and player two within each pair. The unsuccessful trials were categorized into two conditions: *forecasting unsuccessful cooperation* on the next consecutive trial (i.e., failed cooperation trials followed by unsuccessful cooperation) and *forecasting successful cooperation* (i.e., failed cooperation trials followed by successful cooperation on the next consecutive trial). The average number of trials included in the analysis was 46.24 (SD = 14.11) for the condition of forecasting successful cooperation and 48.90 (SD = 26.95) for the condition of forecasting failed cooperation.

To compare the between player coupling of power between the two conditions, we conducted a Spearman's Rho analysis on the power values between player one and player two for every time, frequency, and channel data points within each trial. To account for trial number inconsistencies between the conditions for each pair of participants, the Spearman rho values were converted to Fisher Z scores (Mazaheri, Segaert, et al., 2018). We used non-parametric cluster-based permutation tests to statistically examine condition differences (i.e., *forecasting unsuccessful cooperation* and *forecasting successful cooperation*) in the inter-player coupling of power. This is the same method as described above (in the *Time-Frequency representations of power analysis section),* but, instead of comparing power values, we compare the Fisher Z converted correlation values (between players) at every channel/time/frequency point. These values are the "inter-player power coupling values". We focused our permutation test analyses on the theta (4-7Hz) band (as it has been previously implicated in network changes underlying behavioural adjustment), and we also exploratively analysed the following pre-defined and averaged frequency range bands: alpha (8-14Hz), low beta (15-20Hz), and high beta (20-25Hz). Based on the *feedback-related* results in the Time-Frequency domain, we predefined the following time windows of interest that we used in the permutation tests: 0-0.4sec and 0.6-1.05sec in the theta band, and 0.4-1.65sec in the alpha band. In addition, we exploratively used 0-2sec time windows in the permutation tests. The permutation tests identified electrode clusters that showed the greatest inter-player power coupling differences between conditions *(forecasting unsuccessful cooperation* vs. *forecasting successful cooperation).*

3.5.1.2 Trial-by-trial analysis linking theta dynamics with behavioural outcomes

In addition to categorizing cooperative outcomes in a dichotomous way (by using a pre-defined threshold, categorizing cooperative success within 250ms and failure beyond it – see Methods section for more details), we undertook an analysis that treated the cooperative outcome as a continuous variable. The cooperative outcome as a continuous variable reflected better cooperation with a smaller absolute difference between the RT's of the button presses of player one and two. This allowed us to examine whether the trial-by-trial inter-player theta dynamics are predictive of cooperative outcomes, when the pre-defined categories of success and failure are not present. To do this, we carried out a multiple linear regression model (estimated using Ordinary Least Squares), with the outcome variable being the continuous cooperative outcome (i.e., time difference between the button presses of player 1 and player 2). The predictors were the independent averaged feedback related individual theta power values for player one and player two. The theta power values were averaged over the time window and channels (see Feedback related results section), which showed significant condition differences (successful vs. unsuccessful cooperation) in individual feedback cue locked oscillatory power changes. The proposed model was as follows: cooperative outcome ~ averaged theta power of player one * averaged theta power of player two + absolute theta power difference between players. This was carried out on a trial-by-trial basis. The merit of trial-by-trial analysis is that it can reveal information that would be lost if data were collapsed into a dichotomous mean of a condition (Pernet et al., 2011).

3.5 Results

3.5.1 Participants adhered to task instructions

In order to check whether participants adhered to task instructions of waiting short, medium, and long amounts of time before synchronising their button presses (depending on the auditory cue type), we carried out we carried out a Repeated Measures ANOVA (Greenhouse Geisser adjusted) with the independent variable being the condition type (i.e., short, medium, or long) and the dependent variable being the reaction time. The test revealed that there was a significant effect of condition type on the RT, ($F(1.062, 29.734) = 67.332$, $p < .001$). The RTs significantly varied upon condition in the following way:

High/Short Condition RT < *Medium Condition RT* < *Long/ Low Condition RT*

Post-hoc Least Significant Difference (LSD) pairwise comparisons showed that all the mean differences between each condition was at $p < .001$. The mean RT for the high/ short condition was .544sec (SD = .321), for the medium condition 1.12sec (SD = .587), and for the long/ low condition 1.681 (SD = .951). This means that all the participants adhered to the task instruction and waited the respective lengths of time before pressing their buttons depending on the auditory cue condition.

Additionally, we examined whether the outcome of each trial (i.e., the time difference between the button presses) could be predicted based on the interplay of adjustments made by both players on a behavioural level. To do this, we computed a multiple linear regression model, with the predictor variables being: the behavioural adjustment of player one and behavioural adjustment of player two, as well as their interaction. The outcome variable was the cooperative outcome (i.e., the time difference between the button press of player one and two); cooperative outcome ~ behavioural adjustment of player one * behavioural adjustment of player two. We found that the model explained a statistically significant but weak proportion of variance (F(3, 2647) = 61.40, p < .001), and importantly the interaction of behavioural adjustment of player one and player two was a significant predictor of cooperative outcomes (β = -0.16, p < .001). This suggests that the cooperative outcome is shaped by the dynamic interaction of behavioural adjustments made by both players. See the supplementary material (section: Interplay of behavioural adjustment (between players) was predictive of cooperative outcome) for how behavioural adjustment was calculated and more detailed results).

3.5.2 Feedback-related results

3.5.2.1 Feedback-related oscillatory modulations in the theta band

We compared the TFRs of power between successful and unsuccessful cooperation conditions across all participants (N=59) aligned to the onset of the feedback text. Figure 3.2 displays the TFRs of power locked to the onset of the feedback text for the incorrect, and correct conditions, and the topographic distributions of the condition effect in the theta band. Below we describe the significant oscillatory differences between successful cooperation and unsuccessful cooperation conditions that were identified using non-parametric cluster-based permutation in the theta band (4-7Hz) post feedback text onset. However, we also observed significant condition differences in the feedback related activity (this included evoked and induced activity) in the delta (1-4Hz), alpha (8-14Hz), low beta (15-20Hz) and high beta (20-25Hz) frequency bands as well as ERP differences (see Supplementary Materials and Suppl. Fig. SM.3 for a comprehensive description of all effects).

We observed significant condition differences (*p* = .003) in theta activity maximal over the central and centro-parietal channels. Specifically, theta power was stronger right after the onset of the feedback text in the failed cooperation condition until around 0.4sec compared to the successful cooperation condition.

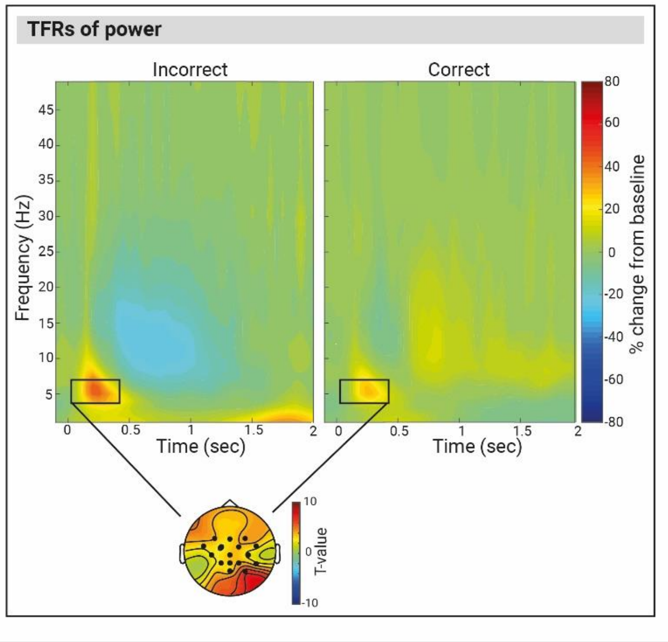


Figure 3.2 Time Frequency representations of power at the Cz electrode for unsuccessful trials (left), and successful trials (right) averaged over all individuals (N=59), locked to the onset of the feedback text. The topoplot illustrates the clusters of electrodes that show the most pronounced condition difference. Black rectangles indicate a significant condition difference (*p* < .05, cluster corrected) in the theta band.

3.5.2.2 Inter-player theta power coupling forecasts cooperative outcome on the next consecutive trial

We compared the inter-player power coupling (see section *Methods: Forecasting cooperative outcome analysis* for how the inter-player power coupling was calculated) between two conditions: *forecasting unsuccessful cooperation* (i.e., failed cooperation trials followed by *a consecutive* unsuccessful cooperation trial) *and forecasting successful cooperation* (i.e., failed cooperation trials followed by *a consecutive* successful cooperation trial) using non-parametric cluster-based permutation tests across all pairs (N = 29 pairs). Figure 3.3 displays inter-player power coupling for (A) *forecasting unsuccessful cooperation* and (B) *forecasting successful cooperation*. We found an inter-player oscillatory power coupling signature that forecasts cooperative outcome on the next consecutive trial. Using the non-parametric cluster-based permutation tests (with a pre-defined time window of 0 - .4sec post feedback based on the feedback-related results in the Time Frequency domain), we observed a significant condition effect (p = .006) in the theta frequency band (this effect was also observed when an exploratory time window of 0 to 2sec was used in the non-parametric cluster-based permutation tests, *p* = .051). This effect corresponded to a cluster that spanned between .15 to .2sec and was maximal over two electrodes (Fz and FC2) that neighboured each other. The inter-player theta power coupling showed the opposite patterns of results in the two conditions. Forecasted cooperative failure was associated with a positive correlation between theta power of player one and player two (i.e. the greater the theta power in one player, the greater the theta power in the other player within a pair or the smaller the theta power in one player, the smaller the theta power in the other player). On the other hand, forecasted cooperative success was linked with a negative correlation between the theta power of player one and player two (i.e., the greater the theta power of one player, the lower the theta power of the other player within a pair or vice versa).

Further, we did not find any significant inter-player power coupling signatures of forecasting cooperative outcomes in the alpha band using the pre-defined .3 to 1.75sec time window. Further analyses using exploratory time window of 0 to 2sec post feedback text also did not yield significant effects in any of the frequency bands (alpha, low beta, and high beta).

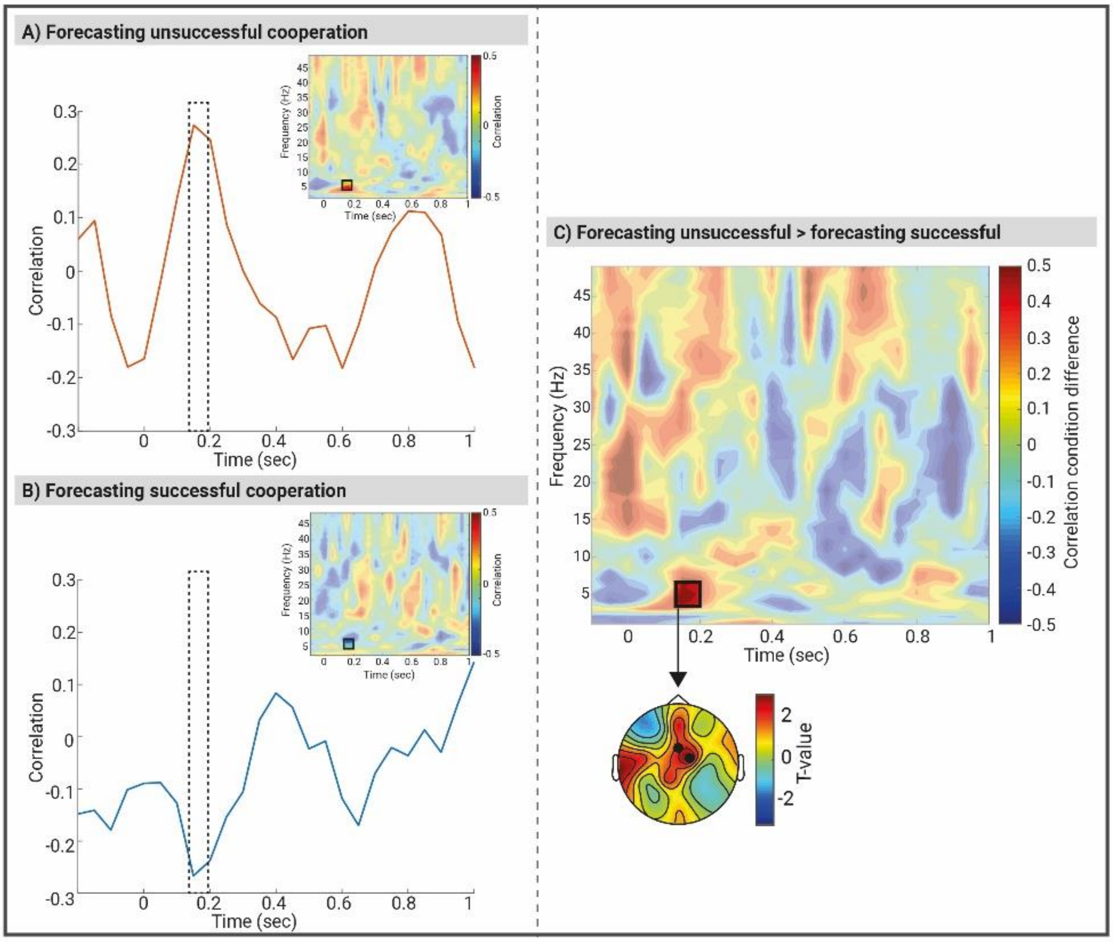


Figure 3.3 Forecasting joint action outcomes (successful vs. unsuccessful cooperation) on next consecutive trials based on the correlation values (Fisher Z) of oscillatory power coupling between player 1 and player 2 after the onset of feedback text in the failed cooperation condition. (A) Line plots are representing the correlation (Fisher Z) values averaged over Fz and FC2 electrodes in the theta frequency band (4-7Hz) forecasting unsuccessful cooperation trials and (B) successful cooperation trials. The dashed rectangles indicate the time windows of the significant between condition differences (as calculated via the non-parametric cluster-based permutation tests). The heatmaps are displaying the correlation values (Fisher Z) of oscillatory power coupling averaged over Fz and FC2 electrodes between player 1 and player 2 locked to the onset of the feedback text forecasting unsuccessful (A) and successful (B) cooperation. The black rectangles indicate significant condition differences (*p* < .05, cluster corrected). (C) Heatmap is representing the condition difference (forecasted unsuccessful cooperation – forecasted successful cooperation) correlation values (Fisher Z) of oscillatory power coupling (averaged across the electrode clusters that indicate maximal condition difference, i.e., Fz and FC2) between player 1 and player 2 locked to the onset of feedback text. The black rectangles indicate significant condition differences (*p* < .0.05, cluster corrected).

3.5.2.3 Relationship between theta dynamics and behaviour on a trial-by-trial basis

All the preceding analyses treated cooperative outcomes as dichotomous variables defined by a pre-determined threshold: cooperative failure vs. success. However, we also sought to examine whether the trial-by-trial inter-player theta dynamics are predictive of continuous cooperative outcomes (i.e., the absolute difference in RT's between player one and player two). We carried out a multiple linear regression analysis as follows: cooperative outcome ~ averaged theta power of player one * averaged theta power of player two + absolute theta power difference between players (see Forecasting cooperative outcome analysis for further details). The model explained a statistically significant and weak proportion of variance (F(4, 2738) = 7.64, p < .001). The model revealed that the averaged theta power of player one (β = 0.08, p < .001), and player two (β = -0.02, p = 0.008) were significant predictors of cooperative outcome. Importantly, the model showed that the interaction of averaged theta power of player one and player two (i.e., averaged theta power of player one * averaged theta power of player two) significantly predicted the continuous cooperative outcome (β = -0.25, p < .001). This suggests that cooperative outcomes are significantly driven by the dynamic interplay of theta power modulations of both players.

Table 3.1 Summary of the multiple linear regression model for the cooperative outcome

Effect	Estimate	SE	*t-value*	*p*
Intercept	.330	.05	6.612	<.001
Theta power of player one	.094	.024	3.845	<.001
Theta power of player two	.061	.023	2.644	.008
Theta power difference (absolute) between players	-.023	.014	-1.675	.094
Theta power of player one * Theta power of player two	-.035	.008	-4.159	<.001

3.6 Discussion

In this study, we investigated whether brain-to-brain coupling patterns in a two-player game could predict performance in a joint-cooperation task. The aim of the task was for the players, upon hearing auditory cues, to establish a shared understanding of ‘short’, ‘medium’, and ‘long’ intervals and synchronise their button presses with each other. The outcome of the joint action was successful only when players converged on their responses. We found that receiving negative feedback led to increased frontal mid-line theta activity across individuals. Moreover, a correlation in post-feedback theta power between players forecasted failed joint action, while an anti-correlation forecasted success. This was evident on a local level (i.e., the inter-player neural dynamics forecasted cooperative outcome on the next consecutive trial, regardless of condition), but not on a global level (i.e., the inter-player dynamics did not forecast the cooperative outcome on the next trial of the same condition type), see Supplementary material “*Effects of brain-to-brain coupling on global cooperative outcomes analysis”* for further details). We suggest that the temporally coupled feedback-related brain activity between two individuals on a local level is an indicator of either redundancy in adjustment of their representation of the task goals or lack of adjustment from both partners. On the other hand, the anti-correlation of this activity reflects cognitive strategic mechanisms that ensure optimal joint action outcomes. The results indicate that successful joint action outcomes are achieved through the implementation of *flexible strategic agility,* whereby in a given trial one partner within the pair demonstrates greater adaptability than the other. This flexible adaptive approach emerges as a key determinant for converging on the representation of the task goals and thus fostering successful cooperation.

Our findings demonstrate how the dynamics of the brain activity induced by feedback between two players could forecast joint action outcomes. With this, we filled a crucial gap in the literature by emphasizing the significance of interpersonal neural patterns in predicting joint action outcomes on subsequent trials by focusing on the correlation of the brain-to-brain neural responses, rather than examining the feedback-related neural responses on an *individual* level.

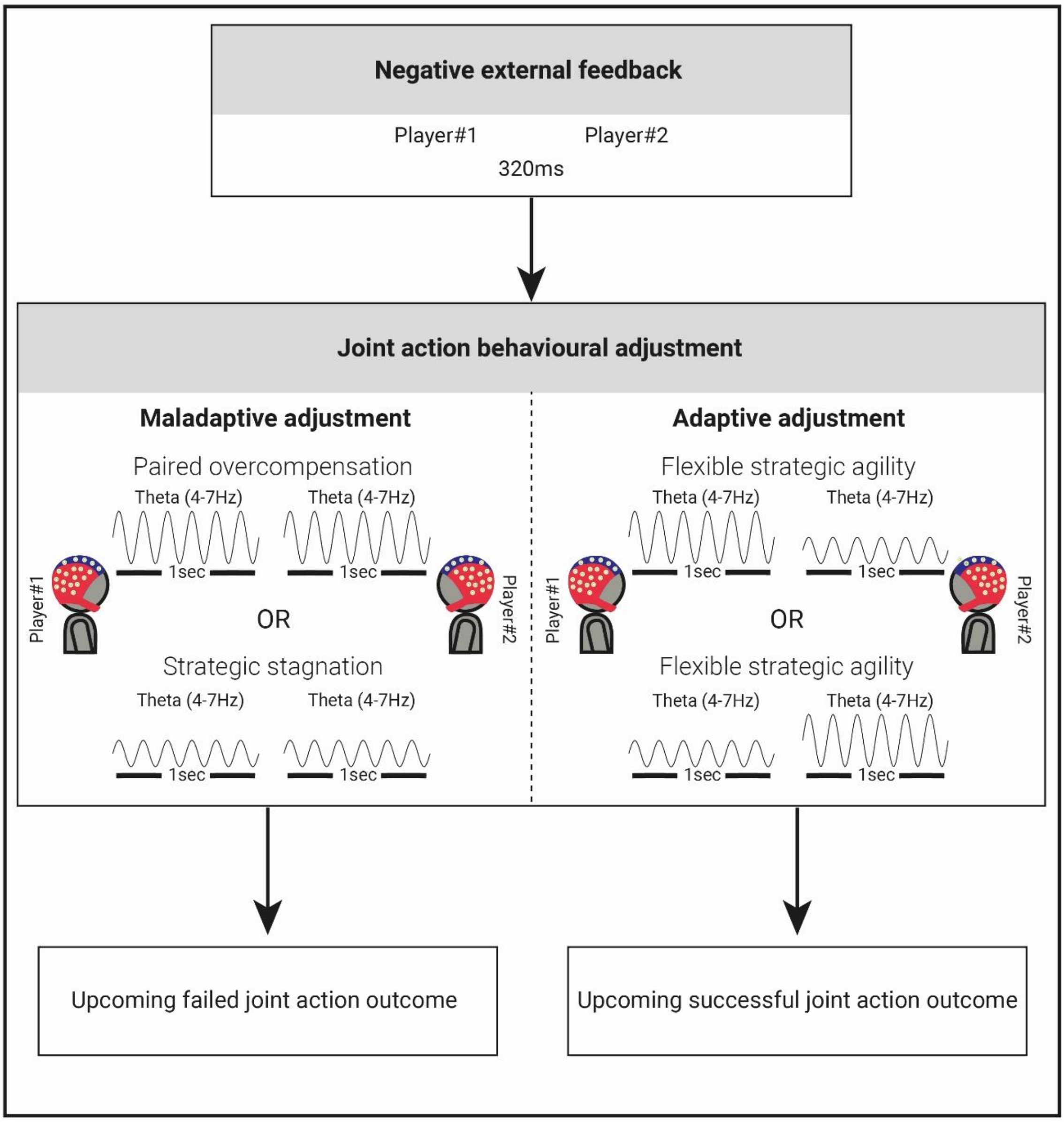


Figure 3.4 Schematic representation of the theta power modulation dynamics between individuals following feedback.

The use of the external feedback allowed the players to adjust their representations of the timing intervals and therefore responses for future trials. On an *individual* level we found distinct patterns of neural activity related to feedback: an increase in frontal midline theta activity after negative feedback and an increase in centro-posterior midline beta following positive feedback. The enhanced theta following negative feedback is linked with enhanced cognitive control and the need for behavioural adjustment following an error, whereas the increased beta following positive feedback signals the requirement for strategy maintenance after a correct response. Most importantly, we found the correlation of brain-to-brain activity in the feedback-related theta power to be a predictor of cooperative outcomes on subsequent trials. Given the well-established role of theta oscillations in adjustment following negative feedback, wherein they override ‘status quo’ responses at individual level, it is reasonable to anticipate distinct brain-to-brain patterns within the theta frequency range, that are predictive of future joint action outcomes at a paired level. Indeed, we found opposite brain-to-brain patterns immediately after the onset of the detailed feedback that forecasted cooperative success and failure. Specifically, *correlated* brain-to-brain activity in the theta power significantly increased the likelihood of an error in the subsequent trial. This can be attributed to two potential cognitive strategies adopted by the players. The two players either passively applied *strategic stagnation*, where neither of the players adjusted their representation of the timing intervals, resulting in failed joint action. Interestingly, a concerted effort of strategy modification from both players (i.e., *paired overcompensation*) in this case resulted in the redundancy of adjustment of the representation of the task goals, overcorrection, and thus subsequent failed joint action. On the other hand, *anti-correlated* brain-to-brain activity in the theta power forecasted successful subsequent trial. Here, players employed a *flexible strategic agility,* wherein one partner adjusted their cognitive strategy/ the representation of the common goal to a greater extent than the other (see Figure 3.4 for a schematic representation of the results). This *flexible strategic agility* facilitated a harmonious balance of adaptive adjustments, resulting in mutual convergence of the time intervals, and thus successful cooperative outcome in the subsequent trial. It is important to note here that while the non-parametric cluster-based permutation tests identified a 50ms time window for this significant predictive effect on cooperative outcomes, this does not imply that the effect was constrained to this specific time frame. It simply means that the effect reached a threshold that indicated significance in this period. However, when visually inspecting Figure 3.3C, it is evident that this effect may be spanning across a larger time window, yet has not met the significance threshold. Hence, we are not suggesting that it only takes 50ms for the participants to adjust their mental representations of the timing intervals and to update their strategies for subsequent trials.

Our findings are consistent with the traditional view of enhanced theta power being linked with enhanced cognitive control (Cavanagh et al., 2009, 2010; Cohen, 2011; van de Vijver et al., 2011). Specifically mid-frontal theta is thought to be a neural marker of the adaptive response to (potential) errors (Cavanagh & Frank, 2014; Cohen, 2014a, 2016). Oscillatory activity in the theta band in the frontal network (including the medial frontal cortex (MFC) and lateral prefrontal cortex (pFC)) is linked to the monitoring of ongoing actions and signalling unfavourable action outcomes. Greater theta power increase has consistently been observed following errors compared to successful outcomes (Cavanagh et al., 2009, 2010; Luu et al., 2004). This theta band activity also extends its role to predicting learning from negative feedback suggesting its involvement in behavioural adjustment in response to feedback that arises from suboptimal actions and thus outcomes (van de Vijver et al., 2011). Specifically, the mPFC theta increase has been linked to predicting post-error reaction time slowing, suggesting greater engagement of cognitive control for identifying and fixing errors (Crivelli-Decker et al., 2018). Conversely, the beta oscillatory power increases have been observed following correct responses and positive feedback (Cohen et al., 2007; Marco-Pallares et al., 2008, 2009) indicating the need for strategy maintenance (Engel & Fries, 2010). This beta band power change facilitates the reinforcement of the current motor or cognitive strategy via the top-down influence (van de Vijver et al., 2011). We replicated and further extended literature findings in this field by demonstrating that modulations in theta and beta band activity in response to feedback are also elicited in joint action settings with the task goal being mutual convergence of timing intervals, rather than individual learned responses.

In addition, our inter-brain findings expanded the traditional view on the link between theta power modulation and cognitive control and moved this framework beyond the potential predictive nature of oscillatory activity on the individual level. Our findings demonstrated that we could apply the current theory to a joint action scenario where its success is reliant upon the convergence and creation of a mutual representation of time intervals between individuals based on the feedback provided. This view is further supported when switching from the dichotomous nature of the joint action outcome (i.e., successful vs. unsuccessful cooperation), and instead examining joint action outcomes as an absolute time difference between players on a trial-by-trial basis. We demonstrated that the interplay of the averaged theta power of player one and player two was a significant predictor of cooperative outcome as measured as a continuous variable (i.e., absolute difference in RT's between the players). The trial-by trial brain dynamics of two interacting players and their subsequent behavioural performance suggests a mechanistic role of the brain signatures we observed and cooperative outcomes.

We revealed distinct strategies that accompanied different joint action outcomes: successful joint action and failed joint action. We demonstrated that flexible joint action adjustment, forecasted future successful joint action. We propose that employing a *flexible strategic agility* allowed for the mutual convergence of shared representations of timing intervals to be developed amongst the players. The *flexible strategic agility* emerged when varying levels of adaptability among partners were developed, with one player exhibiting a greater propensity for strategic adjustment than the other. This could then lead to the harmonious balance of adaptive adjustment, and thus successful cooperation. We posit that this is displayed in the anti-correlated brain-to-brain activity, wherein the higher the theta power in one player, the lower the theta power in the other player. We speculate that this *anti-*

correlated brain-to-brain activity was indicative of the players engaging in rather complementary adjustment processes and increasing the likelihood of uncovering different solutions for convergence on a representation of the common goal.

On the other hand, a correlated brain-to-brain activity in the theta power was foretelling of future failed joint action. We speculate that this could be attributed to the adoption of two distinct cognitive strategies by the players. In some situations, players could have used a passive approach of *strategic stagnation*, wherein both players refrained from adjusting their representation of the timing intervals. The lack of strategic adjustment hindered effective cooperation between the players, impeding their ability to synchronise actions. In other situations, the players could have relied on a *paired overcompensation* strategy with a tendency to follow similar approaches or strategies. This redundancy in adjustment of the common goal could limit the range of perspectives and strategies applied to the task, potentially hindering performance. This flawed dynamic of maladaptive adjustment disrupted the delicate balance required for successful joint action.

Cavanagh et al. (2010) observed post error slowing after negative feedback and speed up following positive feedback of the same stimuli type (after some delay) and not immediately following feedback. They interpreted this effect as indicative of working memory for the specific stimulus type – this is not what we see here, we see a local effect of adjustment reflected in theta power modulations in the players on trials immediately after the feedback. In contrast, we did not observe the same effect for the same condition type stimuli (i.e., the global effect). The difference of this lies potentially in the methodologies, Cavanagh et al. (2010) used a probabilistic learning task in *individual* participants, whereas our participants had to take into account the behavioural adjustment of their partner as well. Post error slowing/ speed up is not applicable in the current study.

Arriving at a shared representation of time through trial-and-error in a joint task can be related to Nash's equilibrium in the context of game theory. Nash's equilibrium is a concept that describes a state in a game where each player's strategy is optimal given the strategies chosen by other players (Jin et al., 2012). It represents a stable outcome where no player has an incentive to unilaterally change their strategy. In the case of arriving at a shared representation of time, the participants in the joint task are engaged in a cooperative endeavour where they need to coordinate their actions based on feedback. Through trial-and-error, they aim to converge on a mutually agreed-upon understanding of 'short', 'medium', and 'long' time intervals. This process involves continuous adjustments and learning from the feedback received. The relation to Nash's equilibrium arises when the participants reach a point where their strategies, in this case, their representations of time intervals, align and are mutually optimal. It signifies that they have found a shared understanding or agreement that maximizes their joint cooperation. This equilibrium state represents a stable and satisfactory outcome in terms of their performance in the joint task.

In the context of our findings, the anti-correlated brain-to-brain activity between individuals suggests a complementary processing style, indicating that each individual is adopting different strategies or approaches. This implies that they are not in a Nash equilibrium because one individual's adjustment in response to feedback is influencing the other individual's subsequent performance. Nash's equilibrium typically involves players making strategic choices based on their beliefs about the other players' actions and seeking to maximize their own outcomes.

As the two-brain science field continues to evolve and researchers increasingly implement neuroscientific hyperscanning methods to study social interaction, it would be important to extend the current findings pertaining the application of feedback in other social scenarios, particularly those that involve verbal communication. By doing so, we could discover valuable insights that can be applied to and improve real-world social interactions such as romantic or friendship relations, team collaborations in workplace settings or even teacher-student interactions in educational settings (e.g., Dikker et al., 2017). On the other hand, we can also delve further into the specific individual characteristics that are associated with the underlying neural processes involved in successful social interactions. For example, as raised by Sebanz et al. (2006) what is the extent to which joint action relies on Theory of Mind (ToM). Building upon our previous work demonstrating a connection between ToM and cooperation (Markiewicz et al., 2023), the next logical step is to explore the relationship between ToM abilities and the inter-person neural correlates of joint action. Furthermore, exploring individual differences in relation to the strategies employed by players and assessing the directionality of the interaction (i.e., leader versus follower) would be an intriguing avenue for future research (e.g., Konvalinka et al., 2010).

In conclusion, we studied how feedback-related changes in the EEG signal between two participants involved in a non-verbal cooperative task adjusted representations of the common goal to ensure optimal joint action outcome. We showed that following external feedback, an anti-correlated brain-to-brain activity in theta power was associated with *flexible strategic agility* and adaptive adjustment leading to successful joint action on subsequent trials. Conversely, a correlated brain-to-brain activity in theta power of the players was linked to either a *strategic stagnation* or *paired overcompensation* strategy; both of which led to maladaptive adjustment and thus failed future joint action. By focusing on the correlation of neural responses and moving beyond an individual-level perspective, we offer novel and more comprehensive insights into how brain-to-brain coupling predicts joint action outcomes.

CHAPTER 4: IT’S NOT ALL ABOUT YOU: COMMUNICATIVE COOPERATION IS DETERMINED BY YOUR PARTNER’S THEORY OF MIND ABILITIES AS WELL AS YOUR OWN

We investigated the relationship between Theory of Mind (ToM) and communicative cooperation. Specifically, we examined whether communicative cooperation is affected by the ToM ability of one’s cooperative partner as well as their own. ToM is the attribution of mental states to oneself and others; cooperation the joint action that leads to achieving a shared goal. We measured cooperation using a novel communicative cooperation game completed by participants in pairs. ToM was measured via the Movies for Assessment of Social Cognition (MASC) task, fluid intelligence via the Raven task. Findings of 350 adults show that ToM scores of both players were predictors of cooperative failure, whereas Raven scores were not. Further, participants were split into low- and high-ToM groups through a median split of the MASC scores: high ToM individuals committed significantly fewer cooperative errors compared to their low ToM counterparts. Therefore, we found a direct relationship between ToM and cooperation. Interestingly, we also examined how ToM scores of paired participants determine cooperation. We found that pairs with two high ToM

individuals committed significantly fewer errors compared to pairs with two low ToM individuals. We speculate that reduced cooperation in low-low ToM pairs is a result of less efficient development of conceptual alignment and recovery from misalignment, compared to high-high ToM dyads. For the first time, we thus demonstrate that it is not all about you; both cooperative partners make key, independent, contributions to cooperative outcomes.

Published in:
Markiewicz, R., Rahman, F., Apperly, I., Mazaheri, A., & Segaert, K. (2023). It is not all about you: Communicative cooperation is determined by your partner's theory of mind abilities as well as your own. Journal of Experimental Psychology: Learning, Memory, and Cognition. https://doi.org/10.1037/xlm0001268

Data availability:
Stimuli, Python scripts for the programmed experiment, and data are available here: https://osf.io/r6p2c/.

4.1 Introduction

The effectiveness of social interactions and cooperation depends on the development of social cognition, which encompasses emotion recognition, empathy, face processing, imitation, and mental state attribution or theory of mind (ToM) (Frith & Frith, 2012). We focus on how ToM affects cooperation between two interacting partners. Players in cooperation games typically share a common goal; it is in their interest to cooperate in order to realise a collective aim (Moll & Tomasello, 2007). Given that cooperation is thought to require predicting, understanding, taking the perspective of, and reasoning about the beliefs and intentions of one's social partner, intuitively, one might expect that individual differences in ToM should relate to cooperative performance. We will empirically investigate this in the present paper.

The drivers of cooperative behaviour have been examined across disciplines such as sociology, primatology and economics. A well-founded account of the (socio-) cognitive mechanisms that drive cooperative behaviour in adult humans, however, remains elusive. To address this, our first study aim was to investigate whether ToM determines communicative cooperation after controlling for fluid intelligence. We measured communicative cooperation using a novel and dynamic symbol matching game, to establish a link between cooperative behaviour in this task and ToM as measured using the Movies for Assessment of Social Cognition paradigm (MASC) (Dziobek et al., 2006). While the MASC is a well-established measure of ToM in the social cognition literature, at the time of writing, there is no existing work that has sought to establish a link between verbal communicative cooperation and ToM (via the MASC) in a healthy young adult population. We further controlled for fluid intelligence. Fluid intelligence has been referred to as logical reasoning and problem solving in novel situations with minimal reliance on current knowledge (Duncan et al., 1995). Previous literature has suggested that fluid intelligence may not only reflect pure cognitive skills but is also related to adaptation to social contexts (Ibanez et al., 2013) and previous studies have shown fluid intelligence to predict ToM (Baker et al., 2014; Ibanez et al., 2013).

Our second aim was to examine for the first time whether verbal communicative cooperation is affected by the ToM competence of one’s cooperative partner in addition to one’s own. We are interested in which pairs (in regard to ToM ability) achieve higher/lower cooperation scores when working together to achieve a shared goal. By using an interactive, real-time, two-player design, we were able to probe more nuanced lines of inquiry such as the effect of having two interactants that are either high-high (both partners have high ToM), low-low (both partners have low ToM), or high-low (one partner has high ToM and the other has low ToM). Do high-high ToM pairs fare better than pairs who are either high-low or low-low? If we find that better *individual* ToM is linked with higher communicative cooperation, then one would expect high-high dyads to demonstrate superior performance in cooperation compared to their low or mixed counterparts. It is worth stressing here that this is an entirely novel approach: there is no existing research that has paired participants of commensurate ToM abilities while seeking to delineate between them with respect to their subsequent cooperation performance. In the cooperation literature, inter-partner ToM is seldom explored; typically, the focus is on differences of *individuals* rather than the collective profile of pairs. Given that social cooperation by its very nature requires human-to-human interaction, it is important to explore how socio-cognitive abilities of pairs drive cooperative behaviour. By contextualising our current understanding of cooperation with ToM, our work has relevance in many real-world settings with important implications for childhood peer interaction, negotiation, social decision making, and even organisational or work-place psychology.

It is important to unpack the term cooperation, since it has been conceptualised in various ways. Cooperative communication games have been used in the psycholinguistics literature since the 1970's to study dialogue (Garrod & Anderson, 1987) and mutual knowledge in conversation (Keysar et al., 2000). Second, cooperation can rely on a non-verbal but mutually salient strategy that leads to mutual understanding (i.e. pure coordination games) (Schelling, 1960). Third, in games such as the Prisoners Dilemma (Schmittberger & Schwarze, 1982) or Ultimatum Game (Poundstone, 1992), cooperation is juxtaposed with defection: if partners decide not to cooperate they are able to negotiate their position. In the present study, we use the term verbal communicative cooperation in a broader sense and bring together aspects of previous literature on both cooperation and communication. Our task does not require interactants to negotiate their position with a view of maximising profit (indeed, there is no reward– monetary or otherwise–for task success) nor do they rely on a salient strategy to form conceptual alignment. In our symbol-matching task, two players are presented with separate pieces of information, which need to be combined via verbal communication. Partners work together with the collective aim of resolving a non-competitive, non-exploitative task and therefore communicative cooperation performance is operationalised as a measure of trial-by-trial error rates. Although the players see different visual information, their roles are not as distinct as in for example Maze Games (Garrod & Anderson, 1987) or the Director's task (Keysar et al., 2000). It is not the case that one participant is the sender of information and the other the mere receiver. There is no structured dialogue but rather a free-flowing conversation with a collective aim of creating mutual understanding leading to clear quantifiable cooperative outcomes. This very much reflects real day-to-day communicative cooperation practices, such as when two people try to move a large heavy sofa up the stairs, which can result in success or failure.

The pure coordination games literature stems from the premise of a focal point (Schelling, 1960): the ability to coordinate without communication by inferring the mutually salient strategy. The most famous example is the New York hypothetical scenario where you and a stranger need to meet – where and when do you go? Surprisingly, there is a consistent intuition to choose the Grand Central station at midday. The salient solution can be helpful in other coordination tasks where communication is not present, like passing someone in the corridor, or passing a junction without clear priority rules. In addition to common knowledge, intuitive alignment supports solving of coordination games without salient answers (Perez-Zapata et al., 2021). Here, we dig deeper into inter-individual differences in conceptual alignment and mutual cooperative behaviour formed via verbal communication.

There is an emerging focus in the cooperation literature on individual differences and other internal drivers of cooperative behaviour. Cooperative behaviour, as measured via the classic Prisoner’s Dilemma Game, has been reported to relate to the Big Five Inventory trait agreeableness (Kagel & McGee, 2014). Previous studies have shown that children as young as six years old are capable of using higher order ToM to coordinate with their peers (Grueneisen et al., 2015). Elsewhere, in a simple two-player “take-it-or-leave-it” Ultimatum game, it has been found that pre-school children who had developed a theory of mind (as measured by a false belief task) suggested a more equitable division of reward (Takagishi et al., 2010). Similarly, ToM predicted performance in a dyadic condition on a spatial, mental rotation task that required children to rely on perspective-taking, false belief understanding, and emotion recognition (Viana et al., 2016). Taken together, these converging lines of evidence suggest that ToM supports cooperative behaviour, albeit mostly in populations of children. Evidence for a link between ToM and cooperation in adults is less clear-cut. There have been reports of better mindreading abilities being linked to greater social cooperation skills (Paal & Bereczkei, 2007), though one should exercise caution: cooperation here was assessed using a self-report questionnaire about personality and behaviour traits of cooperativeness (as opposed to cooperative performance as measured via an ecologically valid task). Conversely, opposite findings were reported when ToM was measured via the Reading the Mind in the Eyes (RMET) task (Baron-Cohen et al., 2001) and cooperation via the classic Prisoner Dilemma Game paradigm: increases in ToM decreased the likelihood of cooperation (DeAngelo & McCannon, 2017). However, it has been recently suggested that the RMET captures predominantly emotion recognition rather than ToM (as is captured by the MASC) (Oakley et al., 2016). Thus, while current literature points toward an association, there is still ongoing debate on the nature and extent to which ToM drives cooperative behaviour. Indeed, more work is needed to evaluate how (if at all) ToM guides cooperation in a healthy, adult sample.

In the current work we used a novel approach to study cooperation, with a task that allows communicative behaviour to unfold as a dynamic process developing over time between two individuals. We posit that the nature of our task is more in keeping with authentic real-world communication (or at least more so than the sometimes contrived classic cooperation games): participants are not privy to their partner’s point of view, they cooperate verbally to reach a shared understanding of a problem, receive regular feedback, and alternate between roles and viewpoints. In real-world communication between humans, we often understand from subsequent actions or communicative responses whether mutual understanding was established. This is a result of reaching conceptual alignment, where ambiguous or novel information is resolved and mental representations of individuals align (Stolk et al., 2016). Much of day-to-day communication and cooperation in society does not depend on posturing one’s position to maximise reward or minimise loss; more often, cooperation is either a mutually beneficial or risk/reward-free enterprise where both parties work collaboratively toward a shared interest.

In the current study we thus examine whether cooperative behaviour that leads to a shared goal relies on ToM. ToM measures that were developed for children typically generate ceiling effects in adults (Apperly & Wang, 2021). The ToM literature has devised different tasks therefore in which mindreading variance is detectable in typical adult populations. For example, the RMET (Baron-Cohen et al., 2001) requires participants to match subtle facial expressions (only showing eyes) to verbal descriptions; the animated triangles task (Castelli et al., 2000) asks participants to describe the behaviour of moving geometric shapes intended to represent social interactions; in the Cambridge Mindreading Face-Voice Battery (Golan et al., 2006), participants are required to select an emotion concept that matches a silent video or voice recording (expressing emotion via facial expression or emotional intonations respectively). For the purpose of understanding and quantifying individual variability in TOM in the current study, we used the MASC (Dziobek et al., 2006), an audio-visual test of mentalizing that depicts social interactions between protagonists in a short film. Periodically, the film is stopped, and participants are asked questions regarding the characters’ beliefs, intentions and actions. The MASC produces an overall mentalizing score as well as sub-scales that evaluate different mentalizing errors (the latter of which we will not use in the present study). The task was developed with a film script and professional actors in order to make the on-screen interactions as life-like as possible (Dziobek et al., 2006). The advantage of the MASC is that it portrays dynamic social scenarios with the rhythm and synchrony of real-world interaction, with actors’ expressions, body language, tonality, and behaviour being considered.

In sum, in the current study we aimed to answer two novel questions. Firstly, we sought to examine the relationship between ToM and communicative cooperation (i.e., verbal communication that leads to a shared goal) when controlling for general abilities (i.e. fluid intelligence). While most of the cited literature suggests that cooperation is indeed driven in part by ToM, the majority of the evidence comes from studies with children and adolescents. A nascent literature suggests that ToM may also be a predictor of cooperative performance in adults. Therefore, we hypothesise a direct link between ToM and communicative cooperation. Specifically, we predict that higher ToM competence will be associated with greater cooperative success and fewer cooperative errors. Secondly, we examined whether cooperation is affected by one's own ToM competence as well as that of one's cooperative partner. Although no research has previously investigated the effect of ToM pairings on cooperative behaviour, intuitively we predict that cooperative performance will be better in the high-high ToM dyads compared to the low-low ToM dyads.

4.2 Materials and Methods

4.2.1 Participants

We recruited 402 healthy participants to take part in our online study (201 participant pairs). The sample size was opportunistic in nature: we selected participants based on their availability and experimenter availability. We excluded 40 participants from the analysis due to: (a) not following instructions in the Communicative cooperation paradigm (resulting in an accuracy below 2SD of the group mean; N=14), (b) internet connection issues during the communicative cooperation paradigm (N=4), (c) not paying attention in part 2 of the study (conducted online without an experimenter present) as indicated by the MASC control questions which probe factual information (score below 2SD of the group mean; N=10), (d) missing data for part 2 of the study (MASC and/or Raven tasks; N=2), (e) not meeting the inclusion criteria of being a fluent speaker in English (N=4), and (f) familiarity of the partner within a pair (i.e., at least one of the participants within a pair reported being either 'friends' or ‘best friends/ partner’; N=8). Above criteria were decided to meet the aim of striking a difficult balance between retaining as much data as possible (since we are interested in co-variance, not in the best estimate of the sample mean) while having to remove some data due to poor quality (online data-collection results in a less controlled environment than an experimental lab, thus we deemed a minimal amount of data removal still necessary). In cases where data of an individual in the pair were removed from the analysis (for any of the reasons outlined above), data of their cooperative partner were unavoidably also removed (N=10). Therefore, the sample analysed consisted of 350 individual participants (175 participant pairs) aged 18 to 34 (M = 19.47, SD = 2.22) (297 females, 49 males, and 4 non-binary individuals).

All participant pairs included reported to not be acquainted with the participant they were paired with. Participants were students at the University of Birmingham and were compensated for their time with course credits or Amazon vouchers. All participants had normal or corrected-to-normal vision, normal hearing, and no neurological or language impairments. All participants included were fluent English speakers; there were 285 native English speakers, of which 207 monolinguals, and 78 spoke at least one other language. Sixty-five participants were not native English speakers but were fluent in English. All participants were given an online information sheet and signed an online consent form prior to taking part in the experiment, which followed the guidelines of the British Psychology Society code of ethics. The experiment was approved by the Science, Technology, Engineering, and Mathematics (STEM) Ethical Review Committee for the University of Birmingham (Ethics Approval Number: ERN_19-1661). Participants were paired at random, relying on the opportunistic sampling strategy.

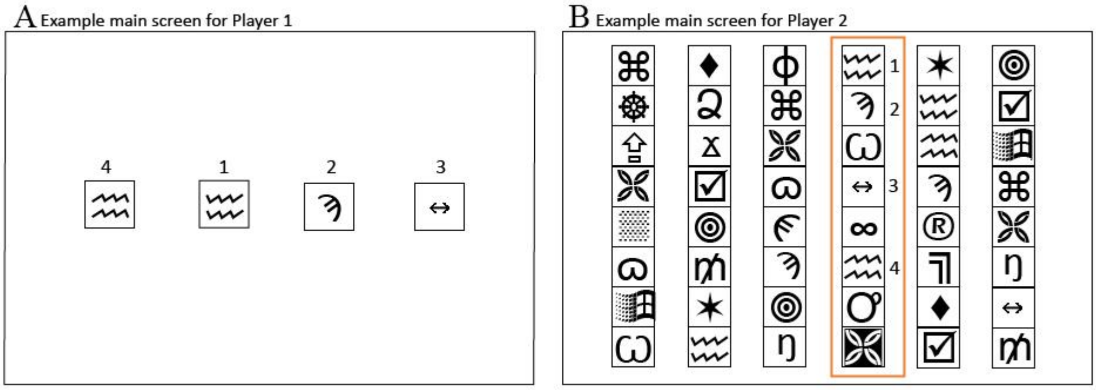


Figure 4.1 (A) Player one first needed to provide a clear description of these four symbols, and then needed to receive information from player two to be able to click on the symbols in the correct order. For the reader's understanding, the numbers above each symbol indicate the order in which the symbols on this trial should be clicked (these were not shown to participants). (B) Based on player one's symbol descriptions, player two had to identify a target column, and then tell player one in which order to click the symbols. For the reader's understanding, the target column is highlighted in orange and the numbers next to symbols reflect the correct order in which player one should click the symbols (again, these elements were not visible to the participant).

4.2.2 Communicative cooperation task

We created a novel experimental task inspired by the game 'Keep talking and nobody explodes' (https://ktane.fandom.com/wiki/Keypad). Each participant was assigned a role, player one or player two, at the beginning of each trial block. Each player was presented with different visual information on their screen: player one could not see the screen of player two and vice versa. The roles switched seven times in total throughout the course of the experiment (after each block). Players had to verbally communicate and cooperate to solve the task, as summarized in Figure 4.1. Note that although players were artificially assigned to 'player one' and 'player two', the task required them to continuously communicate/ cooperate and exchange dialogue just like in a real life scenario. Players were required to combine pieces of different information to solve the task. Player two was presented with 48 symbols organised in 6 columns (such that each column contains 8 symbols) (Figure 4.1B) randomly chosen from a set of 120 symbols with one of the columns being the target column and the rest distractors[1]. At this stage it was unknown to player two which was the target column. Concurrently, player one was presented with 4 symbols[2] randomly chosen from one of the columns (target column) viewed by player two (Figure 4.1A). Player one had to first describe the symbols to player two. Player two was then required to find the correct column that contained all the symbols (amongst other distractor symbols) described by player one. Player two was then to say the order of the symbols in the column. Player one then needed to click on the symbols in the order told by player two. Once a symbol was clicked, a blue border appeared around the symbol (player one could not ‘unclick’ the symbol to correct themselves in case they made a mistake; participants were made aware of this at the start of the experiment as part of the given instructions).

Table 4.1. Two example transcripts of communication between participants within a pair from separate trials. Note that example one matches the visual illustration of the task in Figure 4.1 and 4.2

	Player 1	**Player 2**
Example 1	“Okay, I’d call them waves erm”	
		“Yeah”
	“Yeah waves but like with points going upwards”	
	“And then it’s the same thing but with points going downwards”	
		“Okay, got it”

---

[1] The set of all symbols included 30 single symbols and 45 pairs of symbols. Each symbol in a pair was a slightly altered version of the other (e.g., crescent moon facing to the right or left). Symbols were randomly chosen for the distractor columns from the set of 120. The target column always contained at least one symbol pair (chosen at random) in order to increase difficulty through ambiguity in descriptions. Other symbols in the target column were chosen at random from the set of 120.

[2] The randomly chosen 4 symbols on the player one’s screen could include both, one or none of the pairs of symbols.

	“And then it’s, it’s not dissimilar to a euro sign but it’s pointing towards the bottom left”	
		“Yeah I can see it”
	“And then it’s a double arrow pointing left and right”	
		“Okay, so the first one is waves and to me they look like they are pointing downwards”
	“Okay”	
		“Then it’s the euro symbol, then the arrows and the waves pointing up”
Example 2	“Okay, so the first one, I’d say it’s like two O’s but one is really small and towards the bottom right of the first one. Then it’s two arrows, one pointing up and one pointing down. But there’s a line beneath it”	
		“Ok”
	“Then there’s a box with a tick in it”	
		“Yeah”
	“Then there’s that euro sign from before”	

	“Ok, did you say bottom right with the O?”
“Yes”	
	“Ok, so that’s first”
“Yep”	
	“Then it’s the arrows up and down with the line, then it’s the euro, then it’s the box with a tick”

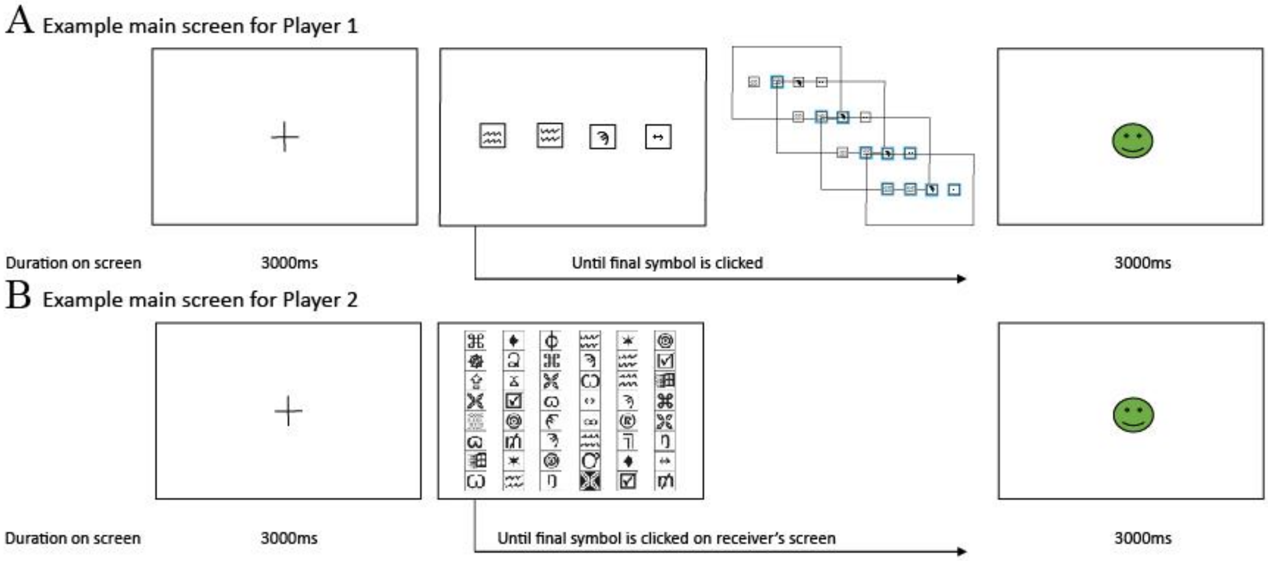


Figure 4.2. Trial presentation and timing for player one (A) and player two (B). The last screen shows the feedback. Successful trials were followed by a green smiley face, unsuccessful trials were by a red sad face, and trials in which the time limit ran out were followed by an orange sad face.

A trial was deemed successful only when all symbols on player one’s screen were clicked in the correct order in the time given for the trial. Upon completion of each trial, both participants received a feedback screen where a green face represented cooperation success, orange represented ‘time ran out’ and red indicated cooperation failure (see Figure 4.2 for timings of each component of a trial and Table 4.1 for representative transcripts of example trials). Each trial had a time limit in which participants were required to respond (i.e. click all four symbols). The time limit for the first 5 trials was 44 seconds. The time limit decreased every 5 trials by 2 seconds with the time limit for the last five trials being 12 seconds. Players were aware there was a time limit but they did not know the length. A red thick line appeared on both screens when the players were close to their time running out (10 seconds). The line decreased every second to indicate remaining time. We manipulated the time limit to increase difficulty and performance variability. Participants were instructed to complete the task as quickly as possible without sacrificing their accuracy.

There were 80 trials in total, divided over 8 blocks (10 trials per block). Each trial led to one of the following outcomes: (1) cooperative success – player one clicked all symbols in the correct order following communication with player two, (2) cooperative failure –player one clicked on all symbols but in the wrong order, or (3) time ran out – player one did not click on all symbols within the time limit. For this study, we focused on cooperative failure and success trials, though for transparency and completeness, we report detailed findings for time ran out trials in the Supplementary Material (see also Suppl. Figure SM.4 and SM.5). We believe that few conclusions can be drawn from time ran out trials: we do not know what the outcome of the ‘time ran out’ trials would have been had participants been given more time (i.e., whether participants were on the right track or not). Time ran out trials are therefore a mixture of ‘almost’ successful and ‘almost’ unsuccessful cooperation. It is very important to keep this in mind when reading the results for these trials in the supplementary materials.

The task was run via Python 3.6 using built in-house and PsychoPy (Peirce, 2007) functions. The experiment scripts and the full set of symbols can be downloaded from https://osf.io/rnzmw/. The task was presented on two identical monitors (one per player) with screen resolution of 1920*1080. Participants remotely accessed the task via TeamViewer or

AnyDesk software. Participants communicated with each other, and the experimenter, via Zoom.

4.2.3 Theory of Mind task

ToM was assessed using the computerised MASC task (Dziobek et al., 2006) administered online via Qualtrics (Qualtrics, Provo, UT). Participants watched 46 short clips (creating one story) and answered questions about the characters' mental states (including their feelings, thoughts and intentions) (e.g., "What is Sandra feeling?"). We embedded 21 control questions. These were simple content questions (e.g., "Which chips does Betty have to play?") to help us determine whether participants paid careful attention to the task.

As we used the multiple choice format of the MASC (Fleck, 2007), each question was scored either as an appropriate or an insufficient ToM response. In previous literature (Hatkevich et al., 2019), insufficient mentalizing responses are sometimes further subdivided into hyper-mentalizing, under-mentalizing, and no mentalizing. Here, we focus only on appropriate mentalizing scores. Higher scores indicated more accurate ToM. As there are 46 experimental questions within the MASC, the minimum possible appropriate mentalizing score is 0 and the maximum is 46.

4.2.4 Fluid Intelligence

The Raven task (Raven, 1958) was used to examine individuals' non-verbal fluid intellectual ability, administered online via Qualtrics (Qualtrics, Provo, UT). The test comprised of 60 (5 sets of 12 items) patterns with a missing section. Participants were asked to indicate (from a series of options) the correct part that fitted the rest of the pattern. The difficulty gradually increased throughout the test. Participants were given 20 minutes to complete it. Correct answers within this time limit were summed.

4.2.5 Procedure

The online experiment consisted of two parts. In part one, we measured cooperation success versus failure in participant pairs. Part two, with the MASC (Dziobek et al., 2006) and Raven (Raven, 1958), was completed individually on a different day.

All participants read an information sheet and signed informed consent online via Qualtrics (Qualtrics, Provo, UT). For part one, two participants joined a Zoom call (see https://osf.io/rnzmw/ for Zoom recording transcripts) with the experimenter, who remained present throughout the communicative cooperation task. This set up is not dissimilar to fMRI hyperscanning studies, in which participants typically communicate via a two-way intercom. Participants accessed the task remotely via TeamViewer or AnyDesk software using a laptop or a PC. Participants were asked to turn their cameras off to avoid gestures conveying symbol information (similar to others restricting the view of dyadic partners (Nadig et al., 2015)). Participants were offered a break in between each block. Upon completion of part one, participants were given access to part two. The ToM and fluid intelligence tasks were administered via Qualtrics (Qualtrics, Provo, UT). Participants completed these individually (without a second participant or experimenter present). There was a compulsory minimum three-minute break between the MASC and the Raven task. Part one of the study lasted approx. one hour, and part two took approx. 45 minutes to complete.

4.3 Data preparation cooperation task

As part one of the study (*cooperation task)* was carried out online (and a Zoom call was a crucial aspect of it), the data collection process was inevitably hindered by internet connection issues. Trials in which the internet connection was momentarily lost by one of the participants or the experimenter were removed (mean number of removed trials per pair = 1.74, SD =

3.27). Further, due to an error in the task programming, some trials suffered from a duplicate symbol, where the same symbol appeared twice. Trials with such instances were also removed (M = 5.21, SD = 2.36). The average number of remaining trials per pair was 73.05 (SD = 3.48).

4.4 Results

4.4.1 High ToM individuals commit significantly fewer cooperative errors compared to low ToM individuals

To examine the effect of individual ToM on cooperation, we conducted a forward multiple regression analysis identifying possible predictors of cooperative failure and success (in separate sets of models), out of the following candidate variables: ToM player one score, ToM player two score, Raven player one score, Raven player two score, a ToM moderator variable (i.e., the interaction between ToM player one and ToM player two score). See Suppl. Table SM.1 for the correlation matrix for outcome and predictor variables. At each step, variables were chosen based on a p-value threshold of <= .05. Data are reported only for variables that remained in the final model with a significance threshold of $p < .05$. The regression model (for cooperative failure) revealed that the ToM player one and ToM player two scores were significant predictors of cooperative failure ($F(2,172) = 4.94$, $p =.008$) and together accounted for 5.4% of variance. More cooperative errors were associated with lower ToM competency in player one ($\beta = -.168$, $p =.025$) and two ($\beta = -.152$, $p = .042$). The interaction between the ToM scores of player one and two was not a significant predictor of cooperative failure ($\beta = -.125$, $p = .095$)., and neither was the Raven score of player one ($\beta = -.085$, $p = .276$) and two ($\beta = .014$, $p = .857$). For cooperative success, the regression model showed that the Raven score of player two accounted for 2.7% ($F(1, 173) = 5.78$, $p = .017$), but the ToM scores did not predict cooperative success.

To visualise above results as well as for the purpose of answering our second research question (see next section), we also report a median split approach. For this, we allocated participants into two categories: participants scoring below or equal to 35 (N=193) were allocated to the ‘low ToM group’ and those scoring above 35 (N=157) were in the ‘high ToM group’. (Those scoring exactly on the median were allocated to the low ToM group to make the groups as equal in size as possible (DeCoster et al., 2011)). The overall range in MASC scores in our sample was 22 – 44. Due to significant differences between the ‘high ToM’ (*M* = 46.255, *SD* = 5.156) and ‘low ToM’ (*M* = 43.653, *SD* = 6.417) groups in Raven scores (*t*(348) = -4.205, *p* < .001) , we controlled for fluid intelligence in the next analyses. A one-way between participants ANCOVA assessed quantitative differences between low vs. high ToM group on cooperative failure and success whilst adjusting for fluid intelligence. There was a significant main effect of ToM group upon cooperative failure, *F*(1, 347) = 11.189, *p* = .001, η^2_p =.031. After adjusting for fluid intelligence, the cooperative failure adjusted mean % for the low ToM group was 13.088 (SEM = .454), and for the high ToM group was 10.79 (SEM = .505) (Fig 4.3 left panel). There was a trend (*F*(1, 347) = 3.172, *p* =.076, η^2_p = .009) suggesting that cooperative success is higher for the high ToM group (adjusted mean % = 48.18, SEM = 1.106) than the low ToM group (adjusted mean % = 45.5, SEM = .995) (Fig 4.3 right panel). This trend mirrors the cooperative failure results.

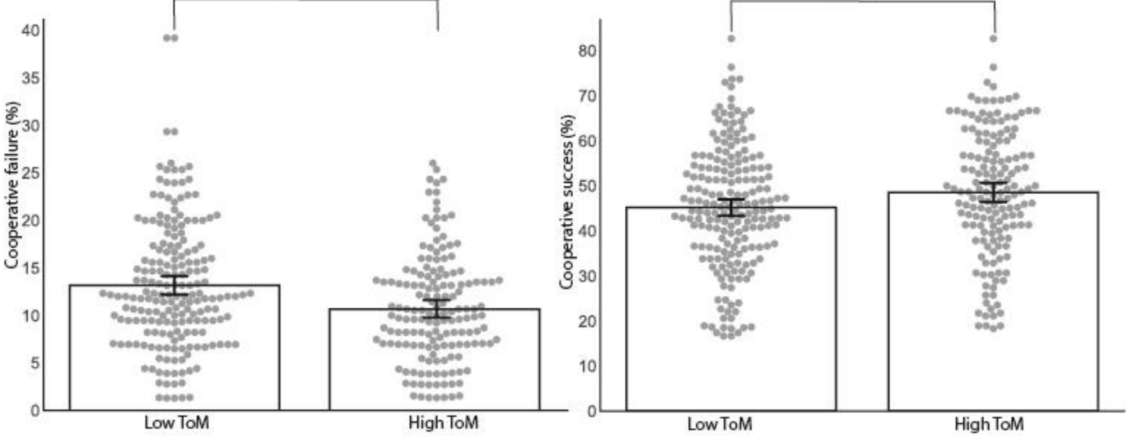


Figure 4.3 Cooperation scores in function of individual theory of mind scores. Bar graphs with mean % score for cooperative failure (left) and cooperative success (right), showing low and high TOM groups. Error bars represent 95% confidence intervals.

4.4.2 Pairs with high-high ToM individuals commit significantly fewer errors when cooperating compared to pairs that consist of low-low ToM individuals

To answer whether cooperation is determined by TOM of *both* partners within a pair, we expand on above and now report the results of a median split analysis with three groups encompassing the ToM scores of both partners within a cooperative pair: (1) low-low ToM pairs (N=57) (i.e. pairs in which both participants within the pair scored ≤35 on the MASC task), (2) mixed ToM pairs (N=79) (i.e. pairs in which one participant within the pair scored ≤35 and the other scored >35 on the MASC task), and (3) high-high ToM pairs (N=39) (i.e. pairs in which both participants within the pair scored >35 on the MASC task). With this approach, we can thus demonstrate how the make-up of a pair of participants determines cooperative failure and success. Examining the effect of participant pairings on cooperative outcome would not be possible without using the initial median split approach above.

We conducted a one-way between participants ANCOVA to quantitatively test for differences between the ToM groups (i.e., low-low, mixed, and high-high) on cooperative failure (Fig 4.4 left panel) and success separately (Fig 4.4 right panel) whilst adjusting for fluid intelligence. In order to account for the impact of fluid intelligence on cooperative failure/ success, we created composite scores of the Raven's test for each pair (i.e., the average Raven's score of the pair) and used this as a covariate in the analysis. There were significant differences between the groups in cooperative failure ($F(2, 172) = 4.968$, $p = .008$, $\eta^2_p = .055$). Further post-hoc Least Significant Difference (LSD) comparisons showed that after adjusting for fluid intelligence there was a significant difference in cooperative failure between the low-low ToM group (adjusted M = 13.971, adjusted SEM =.835) and high-high ToM group (adjusted M = 9.778, adjusted SEM = 1.019) (adjusted mean difference = 4.193, p =.002). The mixed group did not significantly differ in cooperative failure from the other two groups. Further, after adjusting for fluid intelligence, we did not find any group differences in cooperative success (*F*(2, 172) = 1.836, *p* = .163, η^2_p =.021). Figure 4.4 (right panel) depicts that group differences are in the expected direction, but these were non-significant.

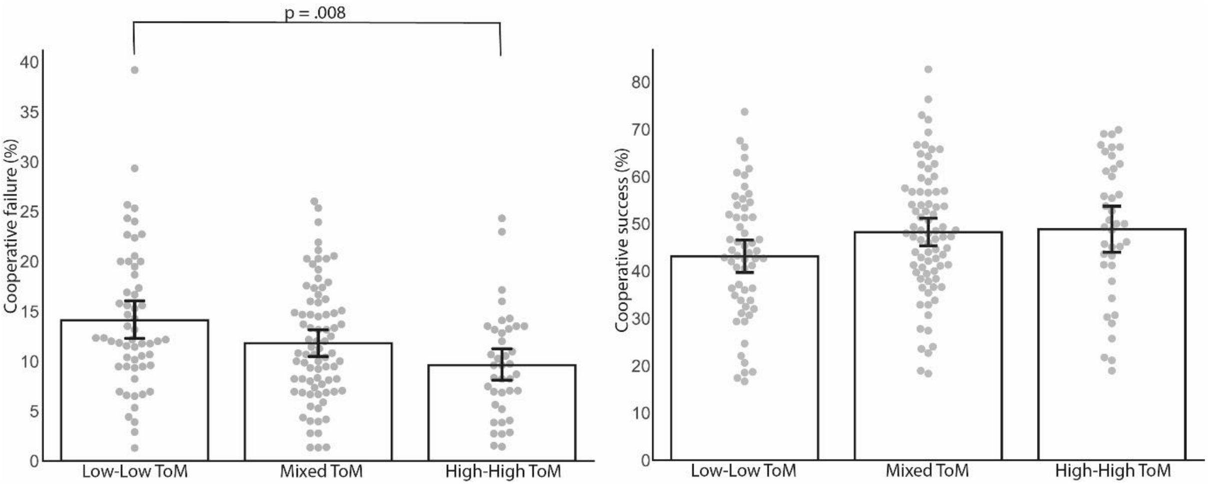


Figure 4.4 Bar graphs for mean % score for cooperative failure (left) and cooperative success (right) showing each group performance separately (low-low, high-high and mixed TOM groups). Error bars represent 95% confidence intervals. Pairs that consisted of high-high ToM individuals committed significantly fewer errors at the communicative cooperation task compared to pairs that consisted of low-low ToM individuals. Although not significant, this pattern is mirrored in the cooperative success scores.

4.5 Discussion

The present study investigated two novel research questions. Firstly, we examined the relationship between ToM and communicative cooperation after controlling for fluid intelligence, and secondly, we investigated whether communicative cooperation is affected by the ToM competence of one's cooperative partner as well as their own. We measured communicative cooperation via a newly developed symbol matching task. ToM was assessed using the MASC (Dziobek et al., 2006) and fluid intelligence via the Raven task (Raven, 1958). Indeed, we found a link between cooperative failure and *individual* ToM scores as measured via the MASC. That is, ToM competency of individual players were significant predictors of cooperative failure, whereas fluid intelligence was not. This was further supported by a median split analysis: after controlling for fluid intelligence, high scoring ToM individuals committed fewer cooperative errors (compared to low scoring ToM individuals). Though merely a trend, the pattern was mirrored in the cooperative success trials. Furthermore, for the first time, we showed that both cooperative partners in the dyad make key, independent, contributions to the cooperative outcome. Dyads with individuals who both scored high on the ToM measure committed fewer cooperative errors compared to pairs of individuals who both scored low on the ToM measure. Again, the pattern was mirrored (but was not significant) in cooperative success trials. The mixed dyads (dyads with one individual scoring high and one individual scoring low on ToM) did not differ from the other dyad types in neither the cooperative failure or success trials.

4.5.1 ToM scores relate to cooperative performance

With respect to the link between *individual* ToM and cooperation, we found that, after controlling for fluid intelligence, individuals who scored high on ToM committed fewer cooperative errors compared to those who scored low on ToM. Analyses with cooperative success as the dependent variable show a trend which mirrors the findings from the cooperative failure trials; that is individuals who scored high on ToM had higher cooperative accuracy compared to those who scored low on ToM (though not significantly so). With this, we have established for the first time with direct and objective measures that there is a relationship between communicative cooperation and ToM in healthy young adults. Our finding is consistent with previous studies that found a link between ToM and cooperation, albeit, using self-reported measures (Paal & Bereczkei, 2007). Moreover, our finding is consistent with literature on a relationship between these two concepts in children (Etel & Slaughter, 2019; Takagishi et al., 2010; Viana et al., 2016).

It is important to note however that although the effects of ToM on cooperative failure were significant, the effect size was small. Hence, we should be cautious with our interpretations. ToM may only be a small contributor to communicative cooperation; other factors may contribute equally, if not more (e.g., religiosity (Xygalatas, 2013), basic personality traits (Thielmann et al., 2014), and generalised reciprocity (Salazar et al., 2022)). On the other hand, measuring ToM is not straightforward: recent attempts to evaluate ToM measures in typical adults show that mindreading task performances do not correlate with one another (Warnell & Redcay, 2019). One possible explanation for this is that laboratory tasks are not effectively assessing real-world relevant abilities but rather are optimised to distinguish between artificial experimental conditions (Apperly & Wang, 2021). It may thus be the case that more sensitive measures of ToM are needed.

A new conversational ToM scale was developed recently, which measures the spontaneous use of ToM during naturalistic conversations using observational ratings for negative (reflecting ToM-related violations of conversational norms) and positive (reflecting mental state language and perspective taking) outcomes (Alkire et al., 2021). Interestingly, the conversational ToM *negative* scale was negatively associated with visuo-affective (assessed via the Cambridge Mindreading Face-Voice Battery for children) (Golan et al., 2006) and spontaneous (assessed via the Triangles task) (Abell et al., 2000) ToM. No association between the conversation ToM *positive* scale and other ToM measures was found. The authors linked the divergence between the negative and positive scales to the multidimensionality of ToM in naturalistic conversation. Individuals who struggle with ToM-related violations in conversation may at the same time display typical levels of other forms of mental state representations as reflected by the conversational ToM positive scale. This relates to our current findings, as the link between ToM and cooperative success was weak (or even absent) compared to the link between ToM and cooperative failure. Cooperative failure and success may be two distinct concepts that tap into different ToM scales. Possibly, the significant difference in cooperative failure between low vs. high ToM individuals reflects conversational violations such as over- or under-informative statements and not explicit references to the partner's mental state.

An important consideration in the study of how ToM and cooperation are inter-linked, is the relationship between ToM and intelligence, as previous studies have shown fluid intelligence to predict ToM. For example, Ibanez et al. (2013) reported that scores on the Raven's progressive matrices were significantly related to performance on Baron-Cohen et al.'s (2001) Reading the Mind in the Eyes Task (or RMET). Further, in a meta-analysis examining the link between IQ— crystallized knowledge and fluid reasoning skills—and performance on the RMET, Baker et al. (2014) reported an overall positive correlation. By considering the contribution of fluid intelligence (as measured via the Raven test (Raven, 1958)), employing a novel, real-time collaborative task alongside a well-established measure of ToM, we identified the link between mental state understanding and cooperative behaviour while controlling for possible contributions of fluid intelligence. Our finding is not dissimilar to the results of Fé et al. (2022), who, using a simple gift-exchange game, found that ToM, but not fluid intelligence, positively predicted intentions-based reciprocity (which helps underpin cooperation).

As previously mentioned, the term *cooperation* in the current study is used in a more collaborative sense than perhaps it was conceptualised in some of the previous literature. Our novel task requires participants to work together towards a collective aim of solving a non-competitive task, instead of working towards maximising profit by negotiating their position (like in for example the Prisoners Dilemma). Similar collaborative games have previously been used to study dialogue (Garrod & Anderson, 1987) and the neural underpinnings of cooperation and competition (Decety et al., 2004; Stolk et al., 2014). However, our current task provides a method for quantifying communicative cooperation with clear-cut outcome measures (i.e., success/failure). Relevant to how we operationalised cooperation is the interactive alignment account (Pickering & Garrod, 2004), which suggests there is automatic alignment of linguistic representations between interlocutors as a result of coupling in production and comprehension, in turn leading to successful communication. It has been argued the communicators develop conceptual alignment in order to resolve ambiguities present in the current communicative signals (Stolk et al., 2016). In relation to our paradigm, the cooperative partners need to develop conceptual alignment when referring to the symbols to create mutual understanding and successful cooperation.

Classical investigations of ToM typically relied on stimuli that were static in nature, with designs that were largely removed from the complexities of social cognition in the real world. In an attempt to better approximate the social cognitive demands of daily life, the MASC requires participants to decode and attribute mental states of characters in a naturalistic film. Since its original publication in 2006, the now well-established MASC has been employed extensively to study social cognition; psychometric evaluation of the instrument has shown it to be a robust and valid measure in both adult clinical and non-clinical samples (Fossati et al., 2018). However, at the time of writing, the MASC has yet to be applied to the study of verbal communicative cooperation with a focus on inter-partner ToM. While previous investigations of ToM and cooperation typically employed tasks and approaches that were less sensitive to socio-cognitive complexity, here, we successfully paired a dynamic and subtle index of ToM with a real-time two-player communicative cooperation task, providing an important step forward for both cooperation and ToM literatures. It is worth mentioning there is evidence showing that variability in performance on ToM tasks may be due to systematic individual differences related to for example depressive symptoms (Nilsen & Duong, 2013). Although not consistent (Ferguson & Cane, 2017), it may be important to control for these factors in future research. In addition, verbal fluency is a worthwhile potential contributor to examine further in this context. Previous research has shown links between verbal fluency and ToM (Ahmed & Miller, 2011). Since our cooperative task relies heavily on language processing, narrative skills may contribute to the relationship between ToM and cooperation as demonstrated in our study. Future research could try to quantify this potential contribution, or, taking a different approach, future studies could instead deploy non-verbal cooperation tasks, where two agents have to decide on action plans together without the language component. Further, the current communicative cooperation task data could be analysed from a different angle. Future analyses could focus on the features of conversation (contained within the conversation transcripts of the experiment) that lead to communication success/ failure, while being linked to ToM competency. This would allow advancing our understanding of the important linguistic mechanisms of communicative success/ failure on an individual level.

4.5.2 Cooperative failure is determined by the ToM of your partner as well as your own

Particularly novel is that we investigated cooperation in relation to ToM of both partners in the dyad. There has been a recent emphasis in the literature on individual differences that govern and affect cooperative behaviour, which has demonstrated links between cooperation and ToM (Paal & Bereczkei, 2007; Takagishi et al., 2010), cultural background (Gächter et al., 2010), and personality traits (Kagel & McGee, 2014). However, it is important to remember that cooperation is the *joint* action of two or more individuals that lead to achieving a shared goal. Therefore, considering only the individual differences of one party involved in cooperative behaviour may be too simplistic. Here, for the first time, we studied the ToM ability of both cooperative partners in order to assess whether the ToM competence of both cooperative partners make independent contributions to cooperative behaviour. In line with our intuitive prediction (note that to date there is no existing literature supporting this line of research), we found that participant pairings that have high-high ToM competence produced significantly fewer cooperative errors compared to participant pairings with low-low ToM abilities. The approach of splitting participants into different ToM competence pairings as well as our novel and real-world-like cooperation paradigm allowed us, for the first time, to show that cooperative outcome is affected by the individual differences of both interlocutors.

The combination of high-high ToM cooperative partners led to superior cooperative performance (compared to low-low ToM partners) in an ecologically valid paradigm. Again, the discrepancy between significant group effects in the cooperative failure vs. success trials can potentially be explained by the proposal that they may reflect positive vs. negative conversational ToM scales (Alkire et al., 2021).

The effect of ToM pairings on cooperative failure can be related to wider theories on verbal interaction, and more particularly, the interactive alignment account (Pickering & Garrod, 2004), which proposes that developing aligned situation models between interlocutors is highly beneficial. In the case of our paradigm, it would be possible for the two cooperative partners to represent the symbols differently (for example, represent the first symbol in Fig.4.1A as either ‘zig zag lines’ or ‘waves’), but it would be inefficient and costly to continuously maintain two different representations of one situation for both partners. We speculate that low-low ToM dyads (so, pairs with partners who both have a low ToM score) might be more likely to represent the symbols differently leading to high computational costs and in turn to poorer performance and more errors. Given that ToM as well as cooperation requires predicting, understanding, and taking the perspective of one’s social partner, we tentatively suggest that dyads that consist of individuals with low-low ToM may have reduced conceptual alignment due to an inefficient development of alignment models compared to those dyads with high-high ToM individuals.

Another important aspect of the interactive alignment account (Pickering & Garrod, 2004) is misalignment and its recovery during communicative exchanges. Misalignment occurs when the representation of the meaning is expressed differently by the communicative partners – for example, the first symbol in Fig. 4.1A (see also example in Table 4.1) could be referred to as “pointing upwards” or “pointing downwards”. Representing the meaning of the symbol in one of the ways over the other may result in communicative misalignment. In this case, an interactive repair is necessary where the partners determine that they cannot simply interpret the input but they must reformulate it to recover from the misalignment. We speculatively suggest that, perhaps, high-high ToM dyads are likely to commit fewer misalignment errors, and when they do, they may recover from them more efficiently as opposed to low-low ToM dyads. Future work including in-depth conversational analyses could shed further light on these issues.

Notably, one of the limitations of our novel paradigm is the time limit. For our time ran out trials, we simply do not know what the outcome would have been if participants had been given an unlimited time. We thus did not focus on these trials in the study. Nonetheless, the inclusion of the time limit may have confounded the measure of cooperation (for example, creating pressure). Then again, if the time limit would have been absent (and participants had been given unlimited time to complete each trial), they may have succeeded in every trial (or at least the majority of them), eliminating individual variability related to cooperative success/ failure. Furthermore, the time limit for each trial decreased gradually throughout the experiment. This was operationalised to increase the difficulty of the task (as it was assumed that cooperative performance would improve over time). Ideally, future studies would incorporate a stable (rather than variable) time limit throughout the whole experiment. This would allow researchers to examine how cooperative behaviour develops over time and how this ties in with ToM competency. Do all pair types (high-high, low-low and mixed ToM pairs) start off at the same cooperation level but the cooperative behaviour improves gradually throughout the task only amongst the high-high ToM pairs and not for the low-low or mixed ToM pairs? Or is it the case that the cooperative performance amongst the high-high ToM pairs is higher (compared to low-low ToM pairs) at the outset and they maintain it throughout, whereas the initial cooperation baseline for the low-low ToM pairs is significantly lower but they gradually increase their cooperative performance as the experiment progresses? One other possibility is that we would see a clear point during the task where the cooperative performance improves significantly for all groups (inferring that there is a set amount of time that individuals need to spend together to develop conceptual alignment). This point may occur significantly later amongst the low-low ToM pairs compared to high-high ToM pairs. Future research can clarify these issues.

Lastly, we would like to contextualise our work in relation to a prominent call in the cognitive neuroscience field with reference to studying the ‘social brain’ and the social-interactive context (Redcay & Warnell, 2018). One way of studying the neural mechanisms that underlie social interaction is by investigating the dynamic relationship between interacting brains (Schoot et al., 2016) using hyperscanning (i.e., measuring brain activity simultaneously from at least two interacting individuals). Investigating the neural underpinnings of social interactions and social decision making has previously been achieved using non-verbal paradigms, with success (Shaw et al., 2018; Stolk et al., 2014). However, as a verbal component to human interaction is often prominent, it seems intuitive that the neural groundworks of verbal communication should be considered also (Salazar et al., 2021). Our current communicative cooperation paradigm would be especially useful for this purpose in future studies. The paradigm offers a clear distinction between successful and unsuccessful verbal cooperation/ communication, therefore neural comparisons (between cooperative success and failure) can easily be drawn. One needs to consider that cooperation/ communication is a joint process, and the cooperating individuals build mutual understanding over time. Using this paradigm in combination with hyperscanning has the potential to pinpoint how the interactants align their behavioural output in temporal and spectral spaces in a dialogue (rather than often used monologue) setting. The paradigm offers a way of studying the neural computational procedures involved in continuous dynamic conceptual alignment and mutual understanding in a real-world-like scenario that leads to clear cut successful or unsuccessful communications.

4.5.3 Implications

A number of training studies have shown that ToM can be improved in children and older adults through carefully constructed training programs (Goldstein & Winner, 2012; Kloo & Perner, 2008; Lecce et al., 2015). Taken together with the current work suggesting that better ToM leads to better cooperation, it is reasonable to postulate that more effective cooperation could result from ToM training. This potentially has important implications in both educational and professional settings. For example, a common pedagogical tool in primary education is group work; here, ToM training may help improve academic and social outcomes for children by way of improved cooperation. Similarly, in the workplace where teamwork is part and parcel of modern working practices, ToM training may lead to improved cooperation.

4.5.4 Conclusion

To summarize, we found a link between communicative cooperation and *individual* ToM competency. Individuals who scored high on the ToM measure committed fewer cooperative errors as opposed to those who scored low on the ToM measure. This is consistent with previous literature on the relationship between these two concepts in young adults as well as children (Etel & Slaughter, 2019; Paal & Bereczkei, 2007; Takagishi et al., 2010; Viana et al., 2016). Most interestingly, we showed for the first time that the ToM competence of both cooperative partners make key, independent, contributions to cooperative failure. Namely, we found that high-high ToM dyads (dyads that consisted of both individuals who scored high on ToM) committed significantly fewer cooperative errors compared to the low-low ToM dyads (dyads that consisted of both individuals who scored low on ToM). Given the requirements for ToM and cooperation, of understanding and predicting the perspective of one’s interactive partner, we postulate that the conceptual alignment in low-low ToM pairs was reduced. We suggest that a reason for this is the inefficient development of alignment models and more frequent misalignment/ less effective recovery from misalignment (Pickering & Garrod, 2004) amongst the low-low ToM compared to high-high ToM pairs.

CHAPTER 5: THESIS GENERAL DISCUSSION

The objective of this thesis was to investigate distinct neural aspects of language and joint action, while also examining how individual differences impact the effectiveness of cooperation. We revealed the neural underpinnings of semantic binding in young adults, and thus established a fundamental reference point against which we could contrast the neural responses of healthy older adults. This approach enabled us to investigate age-related differences in language processing. Through the utilization of two-player paradigms and EEG hyperscanning, this thesis advanced research towards a more ecologically grounded approach to the exploration of human interaction. This progressive step signifies a shift towards a richer understanding of social cognition. The aims of each chapter were:

- To investigate the age-related modulations in oscillatory brain activity related to semantic binding (**Chapter 2**) in a minimal two-word phrase paradigm.

- To examine the brain-to-brain neural patterns that underlie the mechanisms of developing shared representations of a task goal in response to feedback in a non-verbal cooperation task (**Chapter 3**).
- To examine how Theory of Mind abilities influence cooperative success in a verbal cooperation task (**Chapter 4**).

This final chapter will first summarise the key findings, followed by contextualizing these findings within the existing body of literature and their theoretical implications. I will then consider the limitations of the studies, while providing potential avenues for future research.

5.1. Summary of findings

Chapter 2 first examined the underlying oscillatory dynamics of lexical retrieval (recognition of the word form) and semantic binding among young adults in a minimal two-word phrase paradigm. This foundational step provided a reference point, enabling us to compare these neural patterns with those observed amongst healthy older adults. This approach facilitated the identification of age-related neural modulation in language processing. Neural patterns were compared for real words (e.g., swift) vs. letterstrings (e.g., swrfeq) – *lexical retrieval process* and for semantic binding (e.g., horse, preceded by swift) vs. no semantic binding (e.g., horse, preceded by swqrfeq) – *semantic binding* process. In summary older adults exhibited different and delayed (in comparison to young adults) oscillatory signatures for lexical retrieval and most importantly for semantic binding. This highlights that older adults make use of contextual information but how and when this occurs in the brain is different to young adults.

Chapter 3 investigated the brain-to-brain coupling in response to feedback in a two-player non-verbal joint action task and its potential at forecasting future cooperative outcomes. Cooperation was measured via a newly developed time-estimation paradigm, in which participants within pairs had to converge on the meaning of time intervals and synchronise their buttons with each other. On *individual* level we found increased frontal mid-line theta power in response to negative feedback (compared to positive feedback). Most interestingly, we found that a correlation in theta power *between the participants within a pair* was predictive of failed joint action. On the other hand, an anti-correlation of the theta oscillatory patterns between players forecasted successful cooperative outcome.

Chapter 4 examined the relationship between ToM and (verbal) cooperation. Cooperation was measured using a newly developed 2-player cooperation game, in which participants within a dyad verbally communicated to exchange necessary information to solve the task. ToM was measured using the MASC (Dziobek et al., 2006), while also controlling for fluid intelligence abilities using the Raven task (Raven, 1958). The results showed a direct relationship between *individual* ToM abilities and cooperative failure, but not fluid intelligence. More importantly, we found that the pairings of the dyads mattered in regard to their ToM scores. Pairs that contained two high ToM individuals performed better on the cooperation task by committing significantly less errors compared to those pairs that encompassed two individuals with low ToM scores.

5.2. Theoretical implications

5.2.1. Oscillatory mechanisms of lexical retrieval and semantic binding in young adults

The key finding consistent with previous literature and theories is that retrieval of a lexical item was accompanied by a greater theta increase compared to the letter string in young adults. This is not surprising as items charged with more complex structure or meaning require greater resource allocation for their retrieval compared to items with a lesser lexical representation (Bastiaansen et al., 2005; Marinkovic et al., 2012; Mellem et al., 2013). This is the case for visual as well as auditory stimuli (Shahin et al., 2009).

In relation to semantic binding, the second word within the word pair represented the target word. This occurs when lexico-semantic information is retrieved from memory to create a composite meaning encompassing the two-word expression. In the absence of semantic binding, the binding of the two words into a meaningful phrase is not necessary. Among young adults, the process of semantic binding manifested as an anticipatory mechanism surrounding the target (second) word onset. Young adults exhibited a smaller low-beta increase followed by semantic binding compared to the no semantic condition, this effect was centred around the onset of the binding process. As previously proposed, modulations of the beta frequency have been implicated in binding past and present linguistic incoming information (Berghoff et al., 2005; von Stein et al., 1999; Weiss & Mueller, 2012). On the other hand, the direction of the beta modulation is opposite to previous studies (e.g., Lewis et al., 2017; Luo et al., 2010), and reasons for this remain unclear. The timing of this signature points toward a predictive mechanism of semantic binding of items into the wider context.

The primary theoretical significance of this part of the study involving young adults lies in its foundational framework for semantic binding within the time-frequency spectra. This groundwork allows us to draw comparisons with older adults, thereby discerning age-related alterations in language processing.

5.2.2. Delayed and inverse age-related oscillatory mechanisms of lexical retrieval and semantic binding

For *lexical retrieval,* we found distinct and opposite patterns of theta and alpha modulation following the onset of the target word/ letterstring between young and older adults. Young adults elicited a bigger, whereas older adults elicited a smaller increase in the theta power soon after the real word onset (compared to the letterstring). In addition, the healthy older adults elicited the ‘typical’ lexical retrival, similar to that of young adults, (i.e., larger theta increase for real words than letterstrings), ~350ms later than young adults. Speculatively, we suggest that healthy ageing is associated with a delayed retrieval of the lexical item from memory, that is older adults require more time to achieve the same language processing as young adults. Furthermore, the topography of this lexical retrieval signature in the theta power was more widely spread (including bilateral occipital and central regions) as opposed to that of the young adults (left lateralised temporal and parietal electrodes). This aligns with the notion that the aging process may be associated with a reduced engagement of task-specific brain regions, potentially compensated by a broader network involvement when compared to young adults (Cabeza et al., 1997; Grady, 2000). However, it is crucial to acknowledge that these observations remain speculative, as our current data do not provide sufficient topographical evidence for drawing definitive conclusions in this regard. Additionally, there tends to be a reduction in hemispheric asymmetry in older adults (i.e., HAROLD; Cabeza, 2002), while their young counter parts elicit lateralised responses.

Furthermore, the reverse effects of alpha rebound were found between the age groups following the retrieval of the lexical item. The older adults elicited a greater alpha rebound in the lexical retrieval condition compared to the letterstrings, this was not present amongst the young adults. Our interpretation is that older adults lack closure in the processing of the letterstring and they attempt to process the non-lexical item for longer. This is in line with previous work on syntactic binding with pseudo words (Poulisse et al., 2020). This alpha rebound signature must be treated with caution as it occurred just before and around the onset of the second word. Hence it is difficult to disentangle whether this alpha signature maps onto the effects of lexical retrieval or semantic binding. We speculatively propose (although caution should be maintained) that the first part of this alpha signature arose due to the atypical alpha rebound in older adults linked to retrieving of the item, whereas young adults did not show this effect. The part of this alpha signature that spans over and after the second word onset may be implicated in the process of semantic binding, including the anticipatory processing. However, no previous research has reported such effect in the alpha band, therefore we can only propose this as a potential theoretical implication.

In regard to *semantic binding,* we also found age-related oscillatory patterns in the beta band that differed from those in the young adults. Healthy older adults displayed a smaller beta decrease in the semantic binding vs no semantic binding condition approx. 500ms after the onset of the target word. This partially supports the maintenance theory (Lewis & Bastiaansen, 2015), where the no semantic binding condition alerted the system about the requirement for change – this was not necessary for the semantic binding condition. We did not observe this for young adults. We propose that the reason for this is the early ‘anticipatory’ beta power modulation that occurred around the onset of the second word in young adults. The young adults were able to predict the requirement for change during the target word onset. Their system did not need to actively monitor for potential changes, as these had already been addressed during the information intake process rather than afterward. Conversely, older adults had to remain attuned to the system's need for subsequent changes, given their inability to foresee semantic binding during its initial formation.

In summary, this study unveils age-related differences in lexical retrieval and semantic binding processes, shedding light on the temporal dynamics and neural underpinnings of language processing in both young and older adults. These findings establish a foundational reference point, enabling us to contrast healthy aging with less favourable aging trajectories, such as those observed in patients with conditions like mild cognitive impairment or dementia (e.g., Segaert et al., 2022).

5.2.3. Brain-to-brain coupling forecasts joint action convergence in response to feedback

One of the important aspects of joint action is action adjustment as it allows to change the trajectory of own actions in response to the incoming input and the actions of the other. Often, such adjustment is seen as a facilitatory process in cooperative success. However, too much adjustment can also hinder joint action. Based on the brain-to-brain dynamics in response to external feedback in a joint action setting, we propose a new theoretical framework.

As a preliminary validation of our experimental paradigm and in line with an extensive body of research, at the *individual* level we found an enhanced power increase in the theta band over the frontal midline region, triggered in response to negative feedback. This pattern captures the requirement for heightened vigilance, cognitive control, and adaptive responsiveness subsequent to an error occurrence (Cavanagh et al., 2009, 2010; Cohen, 2011; van de Vijver et al., 2011). In addition, the results from the brain-to-brain dynamics have expanded our conventional understanding of the interplay between theta power modulation and cognitive control beyond the *individual* level. We can apply the current framework into a joint action context, where the success of joint action is reliant on the development of convergence and mutual representation of the task goal (i.e., time intervals in the case of the current study) based on the external feedback. We found that a correlation between brain-to-brain activity in the theta power was predictive of failed cooperative outcome in the subsequent trial. Here, we propose two potential strategies that may have been adopted by the partners within the pair. Firstly, it is plausible that both players employed *strategic stagnation*, whereby neither participant revised their representation of the temporal interval, culminating in an unsuccessful button synchrony. Conversely, they might have both opted for a '*meet in the middle*' approach, wherein both individuals adjusted their temporal perceptions. Nevertheless, this simultaneous adjustment could have been excessive, resulting in an overshooting of necessary adaptations. Consequently, this redundancy in representation adjustment likely contributed to the subsequent trial's cooperative failure.

On the other hand, we found that an anti-correlation of theta power brain-to-brain dynamics between the players forecasted successful outcome of the joint action on the subsequent trial. Here, we speculate that the individuals used *flexible strategic agility*, whereby one player adjusted their representation of the timing intervals to a greater extent than the other. The different levels of adaptability resulted in a perfect balance of mutual convergence of shared representation of the temporal intervals and thus subsequent successful joint action. It is crucial to highlight that the time window derived from permutation tests, revealing distinct brain-to-brain patterns predictive of cooperative outcomes, spans a concise 50ms. However, it is important to clarify that we do not attribute the adjustment of mental representations and strategy updates solely to this 50ms timeframe. Rather, it indicates that the effect did not reach significance beyond this timeframe. Refer to Figure 3.3.C for a visual representation, where it becomes evident that this effect extends beyond the 50ms window but falls short of statistical significance.

This research contributes significantly to the existing body of knowledge on feedback mechanisms, highlighting the similarity of oscillatory patterns between feedback responses in joint action contexts and those in individual scenarios. Moreover, these observations rise above individual effects and highlight the importance of considering interpersonal neural patterns in forecasting joint action outcomes. We encompass brain-to-brain dynamics and offer predictive insights into the outcomes of joint actions. The theoretical applications of this study are profound, as we introduce a novel framework that explains how the convergence of task goal representations is reflected through the coupling of two distinct brains.

5.2.4. ToM abilities of both partners are related to their cooperative success

Given the similar underlying mechanisms of ToM and cooperation, that is understanding and predicting the intentions and beliefs of others, it may not be surprising to see a direct relationship between ToM and cooperation on individual level. Specifically, we found that ToM was predictive of cooperation above and beyond fluid intelligence with higher ToM scores being linked with lower cooperative error rates. This pattern of results, although not significant, was mirrored in cooperative success – higher ToM scores were associated with higher cooperative success. Recently, a new ToM scale was conducted, which estimates the spontaneous utilization of ToM during naturalistic dialogues. It was divided into ratings for negative (indicative of breaches in ToM-related conversational norms) and positive outcomes (reflecting the use of mental state language and perspective-taking) (Alkire et al., 2021). This distinction and the multi-layered nature of ToM within conversational contexts may offer insight as to why we observed a relationship between ToM and cooperative failure, but not success. Failed cooperation might be tapping into the facet of ToM that comes into play during naturalistic conversations involving breaches in conversational norms. Conversely, cooperative success could be entwined with the capacity for comprehending mental states and engaging in perspective-taking.

Most importantly and of interest to taking the step toward more ecologically valid research, we found that the ToM ability *pairing* had an effect on the cooperative outcome. Pairs with two individuals with high scores on ToM produced significantly fewer errors in the cooperation task compared to pairings of two low ToM individuals. This verbal cooperation experiment clearly highlights the bilateral nature of dialogue, whereby both, the speaker's and the listener's actions are taken into account in order to develop common grounding (Clark & Krych, 2004).

The relationship between ToM and communicative cooperation is related to the wider theory of social interaction by Pickering and Garrod (2004), the interactive alignment account. This account proposes that in order for successful social interaction to take place, development of aligned situational models and common ground is essential. We speculatively propose that individuals within the low-low ToM pairs represent the stimulus symbols within our paradigm in distinct ways. Consequently, this divergence may hinder the development of common ground and result in less effective alignment models, as opposed to the scenario observed within high-high ToM pairs. Further, in line with the interactive alignment theory, we postulate that the low-low ToM pairs are prone to encountering more frequent instances of misalignment and that their subsequent recovery from such instances might be prolonged in comparison to the high-high ToM pairs.

This study adds to the growing body of research of the intersection of ToM and cooperation. It presents a clear and direct relationship between ToM and communicative cooperation. It highlights the significance of ToM in facilitating shared understanding, alignment, and successful cooperation. A distinctive feature of this study lies in the commitment to examining social interaction in its entirety, wherein the social cognition of both interactive partners is considered. Further, a pivotal avenue for practical implications emerges from this study's findings. If the enhancement of ToM through targeted training programs translates into enhanced cooperation, a logical assumption follows that ToM-focused training might consequently exert a positive influence on cooperative skills as well. This proposal carries profound significance, particularly within educational and professional contexts where collaborative activities such as group work or teamwork hold paramount importance.

5.3. Methodological challenges and future research avenues

5.3.1. Aligning behaviour and neural patterns

Aligned with the central theme of this thesis, which encompasses effective communication, joint action, and cooperation, a clear future step has emerged for **Chapter 2**. While our study advanced the field of language comprehension by investigating the neural foundations of lexical and semantic binding across two distinct populations: young and older adults, it is important to note that our study did not focus on examining the effectiveness of language comprehension. Previous literature has often interpreted different neural mechanisms (between young and healthy older adults) of cognition as compensatory (e.g., Cabeza et al., 2002). However, in order to ascertain whether a neural mechanism is compensatory or not, one needs to consider how the neural patterns contribute to the behavioural performance. Therefore, to assess this, a trial-by-trial examination of the link between neural patterns and behaviour is imperative to test which changes are indeed predictive of successful or effective language processing and in turn communication.

Similarly, while we successfully predicted future joint action outcomes based on brain-to-brain dynamics in **Chapter 3** (i.e., we ascribed this to adjusting and developing convergence of the mental representation of the timing intervals rather than behavioural adjustment itself), we encountered a challenge in establishing a clear link between behaviour (or behavioural adjustment) and neural patterns, which would strengthen the theoretical implications. This limitation stemmed from our methodological approach, which impeded our ability to accurately measure 'behavioural adjustment'. Specifically, we opted to randomly intersperse the various timing interval conditions (high/short, medium, and low/long) within each block, rather than employing a blocked design where a single type of timing interval persists throughout an entire block. We recognised that a blocked design might overly simplify the task due to its predictability, leading to uneven numbers of successful and failed joint action trials. As a result of this methodological aspect, the use of random timing interval distributions confounded the behavioural adjustment variable. To illustrate, if we were to calculate the behavioural adjustment (on subsequent trials) from trial A, *high/short trial*, to trial B, *low/long trial*, the computed behavioural adjustment (RT of trial B subtracted from RT of trial A, for each player separately) might misleadingly appear large. However, this may not accurately reflect the behavioural adjustment influenced by the feedback received; rather, it might predominantly mirror the effects of differing conditions (going from short to long timing intervals). While the potential to compute this behavioural adjustment and establishing a connection with neural responses (or brain-to-brain dynamics exists) exists, such an association would likely yield limited or inconclusive outcomes. As an alternative, future investigations could consider adopting a blocked design that incorporates distinct timing intervals. This would facilitate the accurate calculation of behavioural adjustment, free from confounding variables, consequently enabling a more insightful relationship with neural patterns.

5.3.2. Capturing individual differences

As shown in **Chapter 4**, individual differences, and specifically individual differences in social cognition play a role in joint action or cooperative behaviour outcomes. Therefore, an interesting future avenue would be to consider these individual differences and relate them to the brain-to-brain dynamics in response to joint action and feedback. This would reinforce the theoretical implications highlighted in **Chapter 3**, offering insights into the distinct brain-to-brain coupling patterns exhibited by individuals with varied ToM abilities. Given, that high-high ToM pairs converge quicker on the representation of the stimuli and repair any misalignments more efficiently in a verbal task, this may also translate to a non-verbal task. This would be reflected in the use of the flexible strategic agility strategy (i.e., anti-correlation of theta values between individuals). Further, we would expect more cases (i.e., trials), where strategic stagnation (i.e., lack of adjustment from both partners) or meet in the middle (i.e., redundancy of adjustment) strategies were applied amongst those pairs that contain two low ToM partners, reflected in the correlated brain-to-brain theta power values between the partners.

5.3.3. Exploring the trajectory of cooperation

One specific methodological aspect of the verbal cooperation task (**Chapter 4**) was the implementation of a gradual reduction in the time limit throughout the experiment. This aimed to heighten the task's difficult, with the underlying assumption that cooperative performance would advance as the task unfolded. While this approach proved to be effective for the initial investigation of the relationship between ToM and cooperation, future studies could explore the more nuanced connection between the two. Specifically, future research could employ a consistent time limit across the entire experiment. This would offer an opportunity to build a deeper understanding of the intricate dynamics that shape cooperative behaviour's evolution over time and its interaction with individuals' ToM capacities. This approach holds the potential to distinguish whether ToM pairings predominantly impact the initiation, culmination, or trajectory of cooperative behaviour learning. Expanding upon this, the investigation into the ToM pairings on the trajectory of cooperation could be extended to *non-verbal* cooperation. This could easily be investigated via the use of the already designed non-verbal cooperation task featured in **Chapter 3**, offering a ready means to address it. Originally, this investigation was set to be part of this thesis, but the unforeseen constraints posed by the Covid-19 pandemic prevented its realisation.

5.4. Concluding words

This thesis has offered insights into the diverse facets of social interactions. By employing a synthesis of neuroscience and a two-player approach, this research not only brings us nearer to real-life interactions but also bridges the gap between laboratory settings and the intricacies of genuine human experiences. Firstly, we established a foundational framework of neural mechanisms of lexical retrieval and semantic binding in young adults. In addition, we revealed significant age-related neural modulations in these language comprehension processes. Older adults showed inverse and delayed patterns of processing of binding of semantic structures into two-word phrases, as reflected in the in the distinct modulations of beta power between young and older adults. Additionally, we revealed the brain-to-brain dynamics that drive the establishment of convergence in representing task goals. An anti-correlation of theta power between individuals predicted successful convergence on task goals, while conversely, a correlation of theta power anticipated an unsuccessful outcome in achieving this convergence, and thus joint action. Lastly, we showed the importance of considering individual differences in cooperative outcomes, particularly the pairings of ToM abilities. We found a direct relationship between ToM and cooperation, but most importantly, we revealed that the pairings of ToM influence the success of cooperation. Pairs that contained two high ToM individuals committed significantly fewer cooperative errors and subsequently cooperated better compared to pairs with two low Tom individuals. In summary, this thesis contributes unique research components that enhance our comprehension of human interaction, combining insights from controlled laboratory experiments with a more ecologically grounded approach. By expanding on the methodological aspects presented in this work, future research can delve deeper into the topic of social interaction, leading to more comprehensive understanding of this multifaceted phenomenon.

REFERENCES

Abell, F., Happé, F., & Frith, U. (2000). Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. *Cognitive Development*, *15*(1), 1–16. https://doi.org/10.1016/S0885-2014(00)00014-9

Ahmed, F. S., & Miller, L. S. (2011). Executive function mechanisms of theory of mind. *Journal of Autism and Developmental Disorders*, *41*(5), 667–678. https://doi.org/10.1007/s10803-010-1087-7

Ahn, S., Cho, H., Kwon, M., Kim, K., Kwon, H., Kim, B. S., Chang, W. S., Chang, J. W., & Jun, S. C. (2018). Interbrain phase synchronization during turn-taking verbal interaction—a hyperscanning study using simultaneous EEG/MEG. *Human Brain Mapping*, *39*(1), 171–188. https://doi.org/10.1002/hbm.23834

Alkire, D., McNaughton, K. A., Yarger, H. A., Shariq, D., & Redcay, E. (2021). *Theory of mind in naturalistic conversations between autistic and typically developing children and adolescents*. https://doi.org/10.31234/osf.io/7j2un

Antonenko, D., Brauer, J., Meinzer, M., Fengler, A., Kerti, L., Friederici, A. D., & Flöel, A. (2013). Functional and structural syntax networks in aging. *NeuroImage*, *83*, 513–523. https://doi.org/10.1016/j.neuroimage.2013.07.018

Apperly, I. A., & Wang, J. J. (2021). Mindreading in adults: cognitive basis, motivation, and individual differences. In H. J. Ferguson & E. E. F. Bradford (Eds.), *The Cognitive Basis of Social Interaction Across the Lifespan*. Oxford University Press. https://doi.org/DOI:10.1093/oso/9780198843290.003.0005

Astolfi, L., Toppi, J., Borghini, G., Vecchiato, G., He, E. J., Roy, A., Cincotti, F., Salinari, S., Mattia, D., He, B., & Babiloni, F. (2012). Cortical activity and functional hyperconnectivity by simultaneous EEG recordings from interacting couples of professional pilots. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 4752–4755. https://doi.org/10.1109/EMBC.2012.6347029

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical data base on CD-ROM*.

Babiloni, C., Buffo, P., Vecchio, F., Marzano, N., Del Percio, C., Spada, D., Rossi, S., Bruni, I., Rossini, P. M., & Perani, D. (2012). Brains “in concert”: Frontal oscillatory alpha rhythms and empathy in professional musicians. *NeuroImage*, *60*(1), 105–116. https://doi.org/10.1016/j.neuroimage.2011.12.008

Baggio, G., & Hagoort, P. (2011). The balance between memory and unification in semantics: A dynamic account of the N400. *Language and Cognitive Processes*, *26*(9), 1338–1367. https://doi.org/10.1080/01690965.2010.542671

Baker, C. A., Peterson, E., Pulos, S., & Kirkland, R. A. (2014). Eyes and IQ: A meta-analysis of the relationship between intelligence and “Reading the Mind in the Eyes.” *Intelligence*, *44*(1), 78–92. https://doi.org/10.1016/j.intell.2014.03.001

Balconi, M., Gatti, L., & Vanutelli, M. E. (2018). EEG functional connectivity and brain-to-brain coupling in failing cognitive strategies. *Consciousness and Cognition*, *60*(February), 86–97. https://doi.org/10.1016/j.concog.2018.03.001

Balconi, M., & Vanutelli, M. E. (2017a). Cooperation and competition with hyperscanning methods: Review and future application to emotion domain. *Frontiers in Computational Neuroscience*, *11*(September), 1–6. https://doi.org/10.3389/fncom.2017.00086

Balconi, M., & Vanutelli, M. E. (2017b). When cooperation was efficient or inefficient. Functional near-infrared spectroscopy evidence. *Frontiers in Systems Neuroscience*, *11*(May), 1–10. https://doi.org/10.3389/fnsys.2017.00026

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42*(2), 241–251. https://doi.org/10.1017/S0021963001006643

Bastiaansen, M. C. M., & Hagoort, P. (2006). Oscillatory neuronal dynamics during language comprehension. *Progress in Brain Research*, *159*(06), 179–196. https://doi.org/10.1016/S0079-6123(06)59012-0

Bastiaansen, M. C. M., & Hagoort, P. (2015). Frequency-based segregation of syntactic and semantic unification during online sentence level language comprehension. *Journal of Cognitive Neuroscience*, *27*(11), 2095–2107. https://doi.org/10.1162/jocn

Bastiaansen, M. C. M., Magyari, L., & Hagoort, P. (2010). Syntactic unification operations are reflected in oscillatory dynamics during on-line sentence comprehension. *Journal of Cognitive Neuroscience*, *22*(7), 1333–1347. https://doi.org/10.1162/jocn.2009.21283

Bastiaansen, M. C. M., Oostenveld, R., Jensen, O., & Hagoort, P. (2008). I see what you mean: Theta power increases are involved in the retrieval of lexical semantic information. *Brain and Language*, *106*(1), 15–28. https://doi.org/10.1016/j.bandl.2007.10.006

Bastiaansen, M. C. M., van Berkum, J. J. A., & Hagoort, P. (2002). Event-related theta power increases in the human EEG during online sentence processing. *Neuroscience Letters*, *323*(1), 13–16. https://doi.org/10.1016/S0304-3940(01)02535-6

Bastiaansen, M. C. M., Van Der Linden, M., Ter Keurs, M., Dijkstra, T., & Hagoort, P. (2005). Theta responses are involved in lexical-semantic retrieval during language processing. *Journal of Cognitive Neuroscience*, *17*(3), 530–541. https://doi.org/10.1162/0898929053279469

Beese, C., Vassileiou, B., Friederici, A. D., & Meyer, L. (2019). Age differences in encoding-related alpha power reflect sentence comprehension difficulties. *Frontiers in Aging Neuroscience*, *10*(JUL), 1–14. https://doi.org/10.3389/fnagi.2019.00183

Bemis, D. K., & Pylkkänen, L. (2011). Simple composition: A magnetoencephalography investigation into the comprehension of minimal linguistic phrases. *Journal of Neuroscience*, *31*(8), 2801–2814. https://doi.org/10.1523/JNEUROSCI.5003-10.2011

Bemis, D. K., & Pylkkänen, L. (2013). Basic linguistic composition recruits the left anterior temporal lobe and left angular gyrus during both listening and reading. *Cerebral Cortex*, *23*(8), 1859–1873. https://doi.org/10.1093/cercor/bhs170

Berghoff, C., Weiss, S., & Mueller, H. M. (2005). Processing figurative language: EEG study with coherence analysis. *Proceedings of the 11th Annual Meeting of the Organization for of Human Brain Mapping*, 1576.

Brysbaert, M., Stevens, M., Mandera, P., & Keuleers, E. (2016). How many words do we know? Practical estimates of vocabulary size dependent on word definition, the degree of language input and the participant's age. *Frontiers in Psychology*, *7*(1116), 1–11. https://doi.org/10.3389/fpsyg.2016.01116

Cabeza, R. (2002). Hemispheric asymmetry reduction in older adults: The HAROLD model. *Psychology and Aging*, *17*(1), 85–100. https://doi.org/10.1037/0882-7974.17.1.85

Cabeza, R., Albert, M., Belleville, S., Craik, F. I. M., Duarte, A., Grady, C. L., Lindenberger, U., Nyberg, L., Park, D. C., Reuter-Lorenz, P. A., Rugg, M. D., Steffener, J., & Rajah, M. N. (2018). Maintenance, reserve and compensation: the cognitive neuroscience of healthy ageing. *Nature Reviews Neuroscience*, *19*(11), 701–710. https://doi.org/10.1038/s41583-018-0068-2

Cabeza, R., Anderson, N. D., Locantore, J. K., & McIntosh, A. R. (2002). Aging gracefully: Compensatory brain activity in high-performing older adults. *NeuroImage*, *17*, 1394–

1402. https://doi.org/10.1006/nimg.2002.1280

Cabeza, R., Grady, C. L., Nyberg, L., McIntosh, A. R., Tulving, E., Kapur, S., Jennings, J. M., Houle, S., & Craik, F. I. M. (1997). Age-related differences in neural activity during memory encoding and retrieval: A positron emission tomography study. *Journal of Neuroscience*, *17*(1), 391–400. https://doi.org/10.1523/jneurosci.17-01-00391.1997

Carson, N., Leach, L., & Murphy, K. J. (2018). A re-examination of Montreal Cognitive Assessment (MoCA) cutoff scores. *International Journal of Geriatric Psychiatry*, *33*(2), 379–388. https://doi.org/10.1002/gps.4756

Castelli, F., Happe, F., Frith, U., & Frith, C. (2000). Movement and Mind : A Functional Imaging Study of Perception and Interpretation of Complex Intentional Movement P ... Interpretation of Complex Intentional Movement Patterns. *NeuroImage*, *12*, 314–325. https://doi.org/10.1006/nimg.2000.0612

Cavanagh, J. F., Cohen, M. X., & Allen, J. J. B. (2009). Prelude to and Resolution of an Error: EEG Phase Synchrony Reveals Cognitive Control Dynamics during Action Monitoring. *Journal of Neuroscience*, *29*(1), 98–105. https://doi.org/10.1523/JNEUROSCI.4137-08.2009

Cavanagh, J. F., Eisenberg, I., Guitart-Masip, M., Huys, Q., & Frank, M. J. (2013). Frontal theta overrides Pavlovian learning biases. *Journal of Neuroscience*, *33*(19), 8541–8548. https://doi.org/10.1523/JNEUROSCI.5754-12.2013

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

Cavanagh, J. F., Frank, M. J., Klein, T. J., & Allen, J. J. B. (2010). Frontal theta links prediction errors to behavioral adaptation in reinforcement learning. *NeuroImage*, *49*(4), 3198–3209. https://doi.org/10.1016/j.neuroimage.2009.11.080

Chao, L. L., & Martin, A. (2000). Representation of manipulable man-made objects in the dorsal stream. *NeuroImage*, *12*(4), 478–484. https://doi.org/10.1006/nimg.2000.0635

Chomsky, N. (1995). Language and Nature. *Oxford Journals*, *104*(413), 1–61.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In *Perspectives on Socially Shared Cognition* (pp. 127–149). Washington, DC: APA Books.

Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, *50*, 62–81. https://doi.org/10.1016/j.jml.2003.08.004

Cohen, M. X. (2011). Error-related medial frontal theta activity predicts cingulate-related structural connectivity. *NeuroImage*, *55*(3), 1373–1383. https://doi.org/10.1016/j.neuroimage.2010.12.072

Cohen, M. X. (2014a). A neural microcircuit for cognitive conflict detection and signaling. *Trends in Neurosciences*, *37*(9), 480–490. https://doi.org/10.1016/j.tins.2014.06.004

Cohen, M. X. (2014b). *Analyzing neural time series data: theory and practice*. MIT Press.

Cohen, M. X. (2016). Midfrontal theta tracks action monitoring over multiple interactive time scales. *NeuroImage*, *141*, 262–272. https://doi.org/10.1016/j.neuroimage.2016.07.054

Cohen, M. X., Elger, C. E., & Ranganath, C. (2007). Reward expectation modulates feedback-related negativity and EEG spectra. *NeuroImage*, *35*(2), 968–978. https://doi.org/10.1016/j.neuroimage.2006.11.056

Crivelli-Decker, J., Hsieh, L. T., Clarke, A., & Ranganath, C. (2018). Theta oscillations promote temporal sequence learning. *Neurobiology of Learning and Memory*, *153*(May), 92–103. https://doi.org/10.1016/j.nlm.2018.05.001

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

Curham, K. J., & Allen, J. J. B. (2022). Logic behind EEG frequency analysis. In P. Gable, M. Miller, & E. Bernat (Eds.), *The Oxford Handbook of EEG Frequency* (1st ed., pp.

15–40). Oxford University Press.

Czeszumski, A., Ehinger, B. V., Wahn, B., & König, P. (2019). The social situation affects how we process feedback about our actions. *Frontiers in Psychology*, *10*(361), 1–12. https://doi.org/10.3389/fpsyg.2019.00361

Czeszumski, A., Eustergerling, S., Lang, A., Menrath, D., Gerstenberger, M., Schuberth, S., Schreiber, F., Rendon, Z. Z., & König, P. (2020). Hyperscanning: A Valid Method to Study Neural Inter-brain Underpinnings of Social Interaction. *Frontiers in Human Neuroscience*, *14*(February), 1–17. https://doi.org/10.3389/fnhum.2020.00039

Davidson, D. J., & Indefrey, P. (2007). An inverse relation between event-related and time-frequency violation responses in sentence processing. *Brain Research*, *1158*(1), 81–92. https://doi.org/10.1016/j.brainres.2007.04.082

Davidson, P. R., & Wolpert, D. M. (2003). Motor learning and prediction in a variable environment. *Current Opinion in Neurobiology*, *13*(2), 232–237. https://doi.org/10.1016/S0959-4388(03)00038-2

Davis, S. W., Dennis, N. A., Daselaar, S. M., Fleck, M. S., & Cabeza, R. (2008). Qué PASA? The posterior-anterior shift in aging. *Cerebral Cortex*, *18*(5), 1201–1209. https://doi.org/10.1093/cercor/bhm155

Deacon, D., Dynowska, A., Ritter, W., & Grose-Fifer, J. (2004). Repetition and semantic priming of nonwords: Implications for theories of N400 and word recognition. *Psychophysiology*, *41*(1), 60–74. https://doi.org/10.1111/1469-8986.00120

DeAngelo, G., & McCannon, B. C. (2017). Theory of Mind predicts cooperative behavior. *Economics Letters*, *155*, 1–4. https://doi.org/10.1016/j.econlet.2017.02.009

Decety, J., Jackson, P. L., Sommerville, J. A., Chaminade, T., & Meltzoff, A. N. (2004). The neural bases of cooperation and competition: An fMRI investigation. *NeuroImage*, *23*(2), 744–751. https://doi.org/10.1016/j.neuroimage.2004.05.025

Decety, J., & Sommerville, J. A. (2003). Shared representations between self and other: A social cognitive neuroscience view. *Trends in Cognitive Sciences*, *7*(12), 527–533. https://doi.org/10.1016/j.tics.2003.10.004

DeCoster, J., Gallucci, M., & Iselin, A.-M. R. (2011). Best Practices for Using Median Splits, Artificial Categorization, and their Continuous Alternatives. *Journal of Experimental Psychopathology*, *2*(2), 197–209. https://doi.org/10.5127/jep.008310

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., Rowland, J., Michalareas, G., Van Bavel, J. J., Ding, M., & 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

Dumas, G. (2011). Towards a two-body neuroscience. *Communicative & Integrative Biology*, *4*(3), 349–352. https://doi.org/10.4161/cib.4.3.15110

Dumas, G., Chavez, M., Nadel, J., & Martinerie, J. (2012). Anatomical connectivity influences both intra- and inter-brain synchronizations. *PLoS ONE*, *7*(5). https://doi.org/10.1371/journal.pone.0036414

Dumas, G., Nadel, J., Soussignan, R., Martinerie, J., & Garnero, L. (2010). Inter-brain synchronization during social interaction. *PLoS ONE*, *5*(8). https://doi.org/10.1371/journal.pone.0012166

Dunbar, R. I. M. (1992). Neocortex size as a constraint size in primates on group ecologically. *Journal of Human Evolution*, *20*, 469–493.

Duncan, J., Burgess, P., & Emslie, H. (1995). Fluid intelligence a er frontal lobe lesions. *Neuropsychologia*, *33*(3), 261–268.

Dziobek, I., Fleck, S., Kalbe, E., Rogers, K., Hassenstab, J., Brand, M., Kessler, J., Woike, J.

K., Wolf, O. T., & Convit, A. (2006). Introducing MASC: A movie for the assessment of social cognition. *Journal of Autism and Developmental Disorders*, *36*(5), 623–636. https://doi.org/10.1007/s10803-006-0107-0

Elk, M. Van, Schie, H. T. Van, Zwaan, R. A., & Bekkering, H. (2010). The functional role of motor activation in language processing: Motor cortical oscillations support lexical-semantic retrieval. *NeuroImage*, *50*(2), 665–677. https://doi.org/10.1016/j.neuroimage.2009.12.123

Engel, A. K., & Fries, P. (2010). Beta-band oscillations-signalling the status quo? *Current Opinion in Neurobiology*, *20*(2), 156–165. https://doi.org/10.1016/j.conb.2010.02.015

Etel, E., & Slaughter, V. (2019). Theory of mind and peer cooperation in two play contexts. *Journal of Applied Developmental Psychology*, *60*(October 2017), 87–95. https://doi.org/10.1016/j.appdev.2018.11.004

Fé, E., Gill, D., & Prowse, V. (2022). Cognitive Skills, Strategic Sophistication, and Life Outcomes. *Journal of Political Economy*, *130*(10), 2644–2704. https://doi.org/10.1086/720460

Federmeier, K. D., & Kutas, M. (2005). Aging in context: Age-related changes in context use during language comprehension. *Psychophysiology*, *42*(2), 133–141. https://doi.org/10.1111/j.1469-8986.2005.00274.x

Federmeier, K. D., McLennan, D. B., de Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology*, *39*(2), 133–146. https://doi.org/10.1017/S0048577202001373

Federmeier, K. D., Schwartz, T. J., Van Petten, C., & Kutas, M. (2003). Sounds, Words, Sentences: Age-Related Changes Across Levels of Language Processing. *Psychology and Aging*, *18*(4), 858–872. https://doi.org/10.1037/0882-7974.18.4.858

Ferguson, H. J., & Cane, J. (2017). Tracking the impact of depression in a perspective-taking task. *Scientific Reports*, *7*(14821), 1–9. https://doi.org/10.1038/s41598-017-13922-y

Flanagan, J. R., & Johansson, R. S. (2003). Action plans used in action observation. *Nature*, *424*(6950), 769–771. https://doi.org/10.1038/nature01861

Fleck, S. (2007). *Theory of Mind bei Patienten mit Paranoider Schizophrenie und Patienten mit Asperger-Syndrom: Untersuchungen mit Einem Neuen Filmischen Testverfahren*. Ruhr-University Bochum.

Fossati, A., Borroni, S., Dziobek, I., Fonagy, P., & Somma, A. (2018). Thinking about assessment: Further evidence of the validity of the movie for the assessment of social cognition as a measure of mentalistic abilities. *Psychoanalytic Psychology*, *35*(1), 127–141. https://doi.org/10.1037/pap0000130

Frith, C. D., & Frith, U. (2012). Mechanisms of Social Cognition. *Annual Review of Psychology*, *63*(1), 287–313. https://doi.org/10.1146/annurev-psych-120710-100449

Froehlich, E., Liebig, J., Morawetz, C., Ziegler, J. C., Braun, M., Heekeren, H. R., & Jacobs, A. M. (2018). Same Same But Different: Processing Words in the Aging Brain. *Neuroscience*, *371*, 75–95. https://doi.org/10.1016/j.neuroscience.2017.11.042

Funane, T., Kiguchi, M., Atsumori, H., Sato, H., Kubota, K., & Koizumi, H. (2011). Synchronous activity of two people's prefrontal cortices during a cooperative task measured by simultaneous near-infrared spectroscopy. *Journal of Biomedical Optics*, *16*(7), 077011. https://doi.org/10.1117/1.3602853

Gächter, S., Herrmann, B., & Thöni, C. (2010). Culture and cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1553), 2651–2661. https://doi.org/10.1098/rstb.2010.0135

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*(2), 593–609. https://doi.org/10.1093/brain/119.2.593

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition, 27*(1987), 181–218.

Gastaldon, S., Arcara, G., Navarrete, E., & Peressotti, F. (2020). Commonalities in alpha and beta neural desynchronizations during prediction in language comprehension and production. *Cortex*, *133*, 328–345. https://doi.org/10.1016/j.cortex.2020.09.026

Golan, O., Baron-Cohen, S., & Hill, J. (2006). The Cambridge Mindreading (CAM) Face-Voice Battery: Testing complex emotion recognition in adults with and without Asperger Syndrome. *Journal of Autism and Developmental Disorders*, *36*(2), 169–183. https://doi.org/10.1007/s10803-005-0057-y

Goldstein, T. R., & Winner, E. (2012). Enhancing Empathy and Theory of Mind. *Journal of Cognition and Development*, *13*(1), 19–37. https://doi.org/10.1080/15248372.2011.573514

Grady, C. L. (2000). Functional brain imaging and age-related changes in cognition. *Biological Psychology*, *54*(2000), 259–281. https://doi.org/10.1016/S0301-0511(00)00059-4

Grady, C. L. (2012). The cognitive neuroscience of ageing. *Nature Reviews Neuroscience*, *13*(7), 491–505. https://doi.org/10.1038/nrn3256

Grafton, S. T., Fadiga, L., Arbib, M. A., & Rizzolatti, G. (1997). Premotor cortex activation during observation and naming of familiar tools. *NeuroImage*, *6*(4), 231–236. https://doi.org/10.1006/nimg.1997.0293

Grafton, S. T., & Hamilton, A. F. d. C. (2007). Evidence for a distributed hierarchy of action representation in the brain. *Human Movement Science*, *26*(4), 590–616. https://doi.org/10.1016/j.humov.2007.05.009

Grueneisen, S., Wyman, E., & Tomasello, M. (2015). "I Know You Don't Know I Know..." Children Use Second-Order False-Belief Reasoning for Peer Coordination. *Child Development*, *86*(1), 287–293. https://doi.org/10.1111/cdev.12264

Hagoort, P. (2005). On Broca, brain, and binding: A new framework. *Trends in Cognitive Sciences*, *9*(9), 416–423. https://doi.org/10.1016/j.tics.2005.07.004

Hagoort, P. (2020). The meaning-making mechanism(s) behind the eyes and between the ears. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1791). https://doi.org/10.1098/rstb.2019.0301

Hagoort, P., Baggio, G., & Willems, R. M. (2009). Semantic Unification. In M. S. Gazzaniga (Ed.), *Cognitive Neurosciences* (4th ed, pp. 819–836). MIT Press. https://doi.org/10.1002/int.4550070108

Hagoort, P., Hald, L., Bastiaansen, M. C. M., & Petersson, K. M. (2004). Integration of Word Meaning and World Knowledge in Language Comprehension. *Science*, *304*(5669), 438–441. https://doi.org/10.1126/science.1095455

Hald, L. A., Bastiaansen, M. C. M., & Hagoort, P. (2006). EEG theta and gamma responses to semantic violations in online sentence processing. *Brain and Language*, *96*(1), 90–105. https://doi.org/10.1016/j.bandl.2005.06.007

Hamilton, A. F. d. C. (2021). Hyperscanning: Beyond the Hype. *Neuron*, *109*(3), 404–407. https://doi.org/10.1016/j.neuron.2020.11.008

Hardy, S. M., Segaert, K., & Wheeldon, L. (2020). Healthy Aging and Sentence Production: Disrupted Lexical Access in the Context of Intact Syntactic Planning. *Frontiers in Psychology*, *11*(February). https://doi.org/10.3389/fpsyg.2020.00257

Hari, R., Henriksson, L., Malinen, S., & Parkkonen, L. (2015). Centrality of Social Interaction in Human Brain Function. *Neuron*, *88*(1), 181–193. https://doi.org/10.1016/j.neuron.2015.09.022

Hatkevich, C., Venta, A., & Sharp, C. (2019). Theory of mind and suicide ideation and attempt in adolescent inpatients. *Journal of Affective Disorders*, *256*(May), 17–25. https://doi.org/10.1016/j.jad.2019.05.051

Hermes, D., Miller, K. J., Vansteensel, M. J., Edwards, E., Ferrier, C. H., Bleichner, M. G., van Rijen, P. C., Aarnoutse, E. J., & Ramsey, N. F. (2014). Cortical theta wanes for language. *NeuroImage*, *85*, 738–748. https://doi.org/10.1016/j.neuroimage.2013.07.029

Heyselaar, E., Wheeldon, L., & Segaert, K. (n.d.). Structural priming is supported by different components of non-declarative memory: Evidence from priming across the lifespan. *Journal of Experimental Psychology: Learning, Memory and Cognition*.

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679–709. https://doi.org/10.1037/0033-295X.109.4.679

Ibanez, A., Huepe, D., Gempp, R., Gutiérrez, V., Rivera-Rei, A., & Toledo, M. I. (2013). Empathy, sex and fluid intelligence as predictors of theory of mind. *Personality and Individual Differences*, *54*(5), 616–621. https://doi.org/10.1016/j.paid.2012.11.022

Indefrey, P., & Levelt, W. J. M. (2000). The Neural Correlates of Language Production. In *The new cognitive neurosciences* (2nd ed., pp. 845–865).

Itagaki, S., & Katayama, J. (2008). Self-relevant criteria determine the evaluation of outcomes induced by others. *NeuroReport*, *19*(3), 383–387. https://doi.org/10.1097/WNR.0b013e3282f556e8

Jensen, O., & Mazaheri, A. (2010). Shaping Functional Architecture by Oscillatory Alpha Activity: Gating by Inhibition. *Frontiers in Human Neuroscience*, *4*(186), 1–8. https://doi.org/10.3389/fnhum.2010.00186

Jin, X., Dan, M., Zhang, N., Yu, W., Fu, X., & Das, S. K. (2012). Game theory for infrastructure security: the power of intent-based adversary models. In *Handbook on Securing Cyber-Physical Critical Infrastructure* (pp. 31–53). Elsevier Inc.

Kagel, J., & McGee, P. (2014). Personality and cooperation in finitely repeated prisoner's dilemma games. *Economics Letters*, *124*(2), 274–277. https://doi.org/10.1016/j.econlet.2014.05.034

Kawasaki, M., Yamada, Y., Ushiku, Y., Miyauchi, E., & Yamaguchi, Y. (2013). Inter-brain synchronization during coordination of speech rhythm in human-to-human social interaction. *Scientific Reports*, *3*, 1–8. https://doi.org/10.1038/srep01692

Keenan, E. L. (1979). On surface form and logical form. *Stud. Linguist. Sci*, *8*, 163–203.

Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science*, *11*(1), 32–38. https://doi.org/10.1111/1467-9280.00211

Kielar, A., Deschamps, T., Jokel, R., & Meltzer, J. A. (2018). Abnormal language-related oscillatory responses in primary progressive aphasia. *NeuroImage: Clinical*, *18*(February), 560–574. https://doi.org/10.1016/j.nicl.2018.02.028

Kielar, A., Meltzer, J. A., Moreno, S., Alain, C., & Bialystok, E. (2014). Oscilatory responses to semantic and syntactic violations. *Journal of Cognitive Neuroscience*, *26*(12), 2840–2862. https://doi.org/10.1162/jocn

Kielar, A., Panamsky, L., Links, K. A., & Meltzer, J. A. (2015). Localization of electrophysiological responses to semantic and syntactic anomalies in language comprehension with MEG. *NeuroImage*, *105*, 507–524. https://doi.org/10.1016/j.neuroimage.2014.11.016

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. https://doi.org/10.1016/j.cell.2019.05.022

Kloo, D., & Perner, J. (2008). Training theory of mind and executive control: A tool for improving school achievement? *Mind, Brain, and Education*, *2*(3), 122–127. https://doi.org/10.1111/j.1751-228X.2008.00042.x

Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological Research on Joint Action. Theory and Data. In *Psychology of Learning and Motivation - Advances in Research and Theory* (B.Ross, Vol. 54). Academic Press. https://doi.org/10.1016/B978-0-12-385527-

5.00003-6

Knoblich, G., & Jordan, J. S. (2003). Action Coordination in Groups and Individuals: Learning Anticipatory Control. *Journal of Experimental Psychology: Learning Memory and Cognition*, *29*(5), 1006–1016. https://doi.org/10.1037/0278-7393.29.5.1006

Konvalinka, I., & Roepstorff, A. (2012). The two-brain approach: How can mutually interacting brains teach us something about social interaction? *Frontiers in Human Neuroscience*, *6*(JULY), 1–10. https://doi.org/10.3389/fnhum.2012.00215

Konvalinka, I., Vuust, P., Roepstorff, A., & Frith, C. D. (2010). Follow you, follow me: Continuous mutual prediction and adaptation in joint tapping. *Quarterly Journal of Experimental Psychology*, *63*(11), 2220–2230. https://doi.org/10.1080/17470218.2010.497843

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, *31*(1), 32–59. https://doi.org/10.1080/23273798.2015.1102299

Kutas, M., & Federmeier, K. D. (2011). Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP). *Annual Review of Psychology*, *62*(1), 621–647. https://doi.org/10.1146/annurev.psych.093008.131123

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*, 203–206.

Lam, N. H. L., Schoffelen, J. M., Uddén, J., Hultén, A., & Hagoort, P. (2016). Neural activity during sentence processing as reflected in theta, alpha, beta, and gamma oscillations. *NeuroImage*, *142*, 43–54. https://doi.org/10.1016/j.neuroimage.2016.03.007

Lecce, S., Bottiroli, S., Bianco, F., Rosi, A., & Cavallini, E. (2015). Training older adults on Theory of Mind (ToM): Transfer on metamemory. *Archives of Gerontology and Geriatrics*, *60*(1), 217–226. https://doi.org/10.1016/j.archger.2014.10.001

Lewis, A. G., & Bastiaansen, M. (2015). A predictive coding framework for rapid neural dynamics during sentence-level language comprehension. *Cortex*, *68*, 155–168. https://doi.org/10.1016/j.cortex.2015.02.014

Lewis, A. G., Lemhöfer, K., Schoffelen, J. M., & Schriefers, H. (2016). Gender agreement violations modulate beta oscillatory dynamics during sentence comprehension: A comparison of second language learners and native speakers. *Neuropsychologia*, *89*, 254–272. https://doi.org/10.1016/j.neuropsychologia.2016.06.031

Lewis, A. G., Schoffelen, J. M., Hoffmann, C., Bastiaansen, M., & Schriefers, H. (2017). Discourse-level semantic coherence influences beta oscillatory dynamics and the N400 during sentence comprehension. *Language, Cognition and Neuroscience*, *32*(5), 601–617. https://doi.org/10.1080/23273798.2016.1211300

Lewis, A. G., Wang, L., & Bastiaansen, M. (2015). Fast oscillatory dynamics during language comprehension: Unification versus maintenance and prediction? *Brain and Language*, *148*, 51–63. https://doi.org/10.1016/j.bandl.2015.01.003

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, *74*(6), 431–461. https://doi.org/10.1037/h0020279

Loehr, J. D., Kourtis, D., & Brazil, I. A. (2015). It’s not just my fault: Neural correlates of feedback processing in solo and joint action. *Biological Psychology*, *111*, 1–7. https://doi.org/10.1016/j.biopsycho.2015.08.004

Lucas, H. D., Gupta, R. S., Hubbard, R. J., & Federmeier, K. D. (2019). Adult Age Differences in the Use of Conceptual Combination as an Associative Encoding Strategy. *Frontiers in Human Neuroscience*, *13*(October), 1–17. https://doi.org/10.3389/fnhum.2019.00339

Luo, Y., Zhang, Y., Feng, X., & Zhou, X. (2010). Electroencephalogram oscillations differentiate semantic and prosodic processes during sentence reading. *Neuroscience*, *169*(2), 654–664. https://doi.org/10.1016/j.neuroscience.2010.05.032

Luu, P., Tucker, D. M., & Makeig, S. (2004). Frontal midline theta and the error-related negativity: Neurophysiological mechanisms of action regulation. *Clinical Neurophysiology*, *115*(8), 1821–1835. https://doi.org/10.1016/j.clinph.2004.03.031

Lyu, B., Choi, H. S., Marslen-Wilson, W. D., Clarke, A., Randall, B., & Tyler, L. K. (2019). Neural dynamics of semantic composition. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(42), 21318–21327. https://doi.org/10.1073/pnas.1903402116

Marco-Pallares, J., Cucurell, D., Cunillera, T., Garcia, R., Andres-Pueyo, A., Munte, T. F., & Rodriguez-Fornells, A. (2008). Human oscillatory activity associated to reward processing in a gambling task. *Neuropsychologia*, *46*(1), 241–248. https://doi.org/10.1016/j.neuropsychologia.2007.07.016

Marco-Pallares, J., Cucurell, D., Cunillera, T., Krämer, U. M., Càmara, E., Nager, W., Bauer, P., Schüle, R., Schöls, L., Münte, T. F., & Rodriguez-Fornells, A. (2009). Genetic Variability in the Dopamine System (Dopamine Receptor D4, Catechol-O-Methyltransferase) Modulates Neurophysiological Responses to Gains and Losses. *Biological Psychiatry*, *66*(2), 154–161. https://doi.org/10.1016/j.biopsych.2009.01.006

Marinkovic, K., Rosen, B. Q., Cox, B., & Kovacevic, S. (2012). Event-related theta power during lexical-semantic retrieval and decision conflict is modulated by alcohol intoxication: Anatomically constrained MEG. *Frontiers in Psychology*, *3*(APR), 1–14. https://doi.org/10.3389/fpsyg.2012.00121

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Markiewicz, R., Rahman, F., Apperly, I., Mazaheri, A., & Segaert, K. (2023). It is not all about you: Communicative cooperation is determined by your partner’s theory of mind abilities as well as your own. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. https://doi.org/10.1037/xlm0001268

Markiewicz, R., Segaert, K., & Mazaheri, A. (2021). How the healthy ageing brain supports semantic binding during language comprehension. *European Journal of Neuroscience*, *54*(11), 7899–7917. https://doi.org/10.1111/ejn.15525

Marsh, K. L., Richardson, M. J., Baron, R. M., & Schmidt, R. C. (2006). Contrasting approaches to perceiving and acting with others. *Ecological Psychology*, *18*(1), 1–38. https://doi.org/10.1207/s15326969eco1801_1

Mazaheri, A. (2022). The relationship between evoked and induced EEG/MEG changes. In P. Gable, M. Miller, & E. Bernat (Eds.), *The Oxford Handbook of EEG Frequency* (1st ed., pp. 115–130). Oxford University Press.

Mazaheri, A., Nieuwenhuis, I. L. C., Van Dijk, H., & Jensen, O. (2009). Prestimulus alpha and mu activity predicts failure to inhibit motor responses. *Human Brain Mapping*, *30*(6), 1791–1800. https://doi.org/10.1002/hbm.20763

Mazaheri, A., & Picton, T. W. (2005). EEG spectral dynamics during discrimination of auditory and visual targets. *Cognitive Brain Research*, *24*(1), 81–96. https://doi.org/10.1016/j.cogbrainres.2004.12.013

Mazaheri, A., Segaert, K., Olichney, J., Yang, J. C., Niu, Y. Q., Shapiro, K., & Bowman, H. (2018). EEG oscillations during word processing predict MCI conversion to Alzheimer’s disease. *NeuroImage: Clinical*, *17*(June 2017), 188–197. https://doi.org/10.1016/j.nicl.2017.10.009

Mazaheri, A., Slagter, H. A., Thut, G., & Foxe, J. J. (2018). Orchestration of brain oscillations: principles and functions. *European Journal of Neuroscience*, *48*(7), 2385–2388. https://doi.org/10.1111/ejn.14189

Mellem, M. S., Bastiaansen, M. C. M., Pilgrim, L. K., Medvedev, A. V., & Friedman, R. B.

(2012). Word class and context affect alpha-band oscillatory dynamics in an older population. *Frontiers in Psychology, 3*(97), 1–8. https://doi.org/10.3389/fpsyg.2012.00097

Mellem, M. S., Friedman, R. B., & Medvedev, A. V. (2013). Gamma- and theta-band synchronization during semantic priming reflect local and long-range lexical-semantic networks. *Brain and Language*, *127*(3), 440–451. https://doi.org/10.1016/j.bandl.2013.09.003

Meltzer, J. A., Kielar, A., Panamsky, L., Links, K. A., Deschamps, T., & Leigh, R. C. (2017). Electrophysiological signatures of phonological and semantic maintenance in sentence repetition. *NeuroImage*, *156*, 302–314. https://doi.org/10.1016/j.neuroimage.2017.05.030

Menenti, L., Gierhan, S. M. E., Segaert, K., & Hagoort, P. (2011). Shared language: Overlap and segregation of the neuronal infrastructure for speaking and listening revealed by functional MRI. *Psychological Science*, *22*(9), 1173–1182. https://doi.org/10.1177/0956797611418347

Meyer, L. (2018). The neural oscillations of speech processing and language comprehension: state of the art and emerging mechanisms. *European Journal of Neuroscience*, *48*(7), 2609–2621. https://doi.org/10.1111/ejn.13748

Meyer, L., Obleser, J., & Friederici, A. D. (2013). Left parietal alpha enhancement during working memory-intensive sentence processing. *CORTEX*, *49*(3), 711–721. https://doi.org/10.1016/j.cortex.2012.03.006

Moll, H., & Tomasello, M. (2007). Cooperation and human cognition: the Vygotskian intelligence hypothesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1480), 639–648. https://doi.org/10.1098/rstb.2006.2000

Montague, P. R., Berns, G. S., Cohen, J. D., McClure, S. M., Pagnoni, G., Dhamala, M., Wiest, M. C., Karpov, I., King, R. D., Apple, N., & Fisher, R. E. (2002). Hyperscanning: Simultaneous fMRI during linked social interactions. *NeuroImage*, *16*(4), 1159–1164. https://doi.org/10.1006/nimg.2002.1150

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

Nadig, A., Seth, S., & Sasson, M. (2015). Global similarities and multifaceted differences in the production of partner-specific referential pacts by adults with autism spectrum disorders. *Frontiers in Psychology, 6*(1888), 1–14. https://doi.org/10.3389/fpsyg.2015.01888

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Nelson, H. E. (1982). The National Adult Reading Test (NART): Test Manual. *Windsor, UK: NFER-Nelson*, *124*(3), 0–25. https://doi.org/Thesis_references-Converted #319

Nilsen, E. S., & Duong, D. (2013). Depressive symptoms and use of perspective taking within a communicative context Depressive symptoms and use of perspective taking within a communicative context. *Cognition & Emotion*, *27*(2), 335–344. https://doi.org/10.1080/02699931.2012.708648

Oakley, B. F. M., Brewer, R., Bird, G., & Catmur, C. (2016). Theory of mind is not theory of emotion: A cautionary note on the reading the mind in the eyes test. *Journal of Abnormal Psychology*, *125*(6), 818–823. https://doi.org/10.1037/abn0000182

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience, 2011*, 1–9. https://doi.org/10.1155/2011/156869

Paal, T., & Bereczkei, T. (2007). Adult theory of mind, cooperation, Machiavellianism: The effect of mindreading on social relations. *Personality and Individual Differences*, *43*(3), 541–551. https://doi.org/10.1016/j.paid.2006.12.021

Peelle, J. E. (2019). Language and Aging. In *The Oxford Handbook of Neurolinguistics* (pp. 1–21). Oxford University Press. https://doi.org/10.1017/s0267190500001203

Peirce, J. W. (2007). PsychoPy-Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Perez-Zapata, D. I., McKenzie-Smart, X., Charest, I., & Apperly, I. (2021). *Meeting and Missing Minds: Children and Adults Use Alignment of Intuitions to Solve Pure Coordination Games*. *July 3*. https://doi.org/10.31234/osf.io/96hbp

Pernet, C. R., Sajda, P., & Rousselet, G. A. (2011). Single-trial analyses: Why bother? *Frontiers in Psychology*, *2*(322), 1–2. https://doi.org/10.3389/fpsyg.2011.00322

Pickering, M. J., & Garrod, S. (2004). The interactive-alignment model: Developments and refinements. *Behavioral and Brain Sciences*, *27*(2), 212–225. https://doi.org/10.1017/s0140525x04450055

Picton, L., Saunders, B., & Jentzsch, I. (2012). "I will fix only my own mistakes" : An ERP study investigating error processing in a joint choice-RT task. *Neuropsychologia*, *50*(5), 777–785. https://doi.org/10.1016/j.neuropsychologia.2012.01.011

Poulisse, C., Wheeldon, L., Limachya, R., Mazaheri, A., & Segaert, K. (2020). The oscillatory mechanisms associated with syntactic binding in healthy ageing. *Manuscript Submitted for Publication*.

Poulisse, C., Wheeldon, L., & Segaert, K. (2019). Evidence against preserved syntactic comprehension in healthy ageing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(12), 2290–2308. https://doi.org/https://doi.org/10.1037/xlm0000707

Poundstone, W. (1992). *Prisoner's Dilemma*. New York: Doubleday.

Prystauka, Y., & Lewis, A. G. (2019). The power of neural oscillations to inform sentence comprehension: A linguistic perspective. *Language and Linguistics Compass*, *13*(9), 1–40. https://doi.org/10.1111/lnc3.12347

Pylkkanen, L. (2019). The neural basis of combinatory syntax and semantics. *Science*, *366*(6461), 62–66. https://doi.org/10.1126/science.aax0050

Pylkkänen, L., Bemis, D. K., & Elorrieta, E. B. (2014). Building phrases in language production: An MEG study of simple composition. *Cognition*, *133*(2), 371–384. https://doi.org/10.1016/j.cognition.2014.07.001

Raven, J. C. (1958). *Standard Progressive Matrices. Sets A, B, C, D, and E.* London: H.K. Lewis & Co. Ltd.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422. https://doi.org/10.1037//0033-2909.124.3.372

Rayner, K., & Clifton, C. (2009). Language Processing in Reading and Speech Perception is Fast and Incremental. *Biol Psychol*, *80*(1), 4–9. https://doi.org/10.1016/j.biopsycho.2008.05.002.Language

Redcay, E., & Warnell, K. R. (2018). A Social-Interactive Neuroscience Approach to Understanding the Developing Brain. In *Advances in Child Development and Behavior* (1st ed., Vol. 54). Elsevier Inc. https://doi.org/10.1016/bs.acdb.2017.10.001

Rommers, J., Dickson, D. S., Norton, J. J. S., Wlotko, E. W., & Federmeier, K. D. (2017). Alpha and theta band dynamics related to sentential constraint and word expectancy. *Language, Cognition and Neuroscience*, *32*(5), 576–589. https://doi.org/10.1080/23273798.2016.1183799

Salazar, M., Joel Shaw, D., Czekóová, K., Staněk, R., & Brázdil, M. (2022). The role of generalised reciprocity and reciprocal tendencies in the emergence of cooperative group norms. *Journal of Economic Psychology*, *90*(102520), 1–13. https://doi.org/10.1016/j.joep.2022.102520

Salazar, M., Shaw, D. J., Gajdoš, M., Mareček, R., Czekóová, K., Mikl, M., & Brázdil, M. (2021). You took the words right out of my mouth: Dual-fMRI reveals intra- and inter-personal neural processes supporting verbal interaction. *NeuroImage*, *228*(117697), 1–11. https://doi.org/10.1016/j.neuroimage.2020.117697

Salthouse, T. A. (1996). The Processing-Speed Theory of Adult Age Differences in Cognition. *Psychological Review*, *103*(3), 403–428. https://doi.org/10.1037/0033-295X.103.3.403

Salthouse, T. A. (2010). Selective review of cognitive aging. *Journal of the International Neuropsychological Society*, *16*(5), 754–760. https://doi.org/10.1017/S1355617710000706

Sänger, J., Müller, V., & Lindenberger, U. (2012). Intra- and interbrain synchronization and network properties when playing guitar in duets. *Frontiers in Human Neuroscience*, *6*(NOVEMBER 2012), 1–19. https://doi.org/10.3389/fnhum.2012.00312

Schelling, T. (1960). *The strategy of conflict*. Harvard University Press.

Schmittberger, G. W., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization*, *3*, 367–388.

Schoot, L., Hagoort, P., & Segaert, K. (2016). What can we learn from a two-brain approach to verbal interaction? *Neuroscience and Biobehavioral Reviews*, *68*, 454–459. https://doi.org/10.1016/j.neubiorev.2016.06.009

Schurz, M., Kronbichler, M., Crone, J., Richlan, F., Klackl, J., & Wimmer, H. (2014). Top-down and bottom-up influences on the left ventral occipito-temporal cortex during visual word recognition: An analysis of effective connectivity. *Human Brain Mapping*, *35*(4), 1668–1680. https://doi.org/10.1002/hbm.22281

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences*, *10*(2), 70–76. https://doi.org/10.1016/j.tics.2005.12.009

Sebanz, N., & Knoblich, G. (2009). Prediction in Joint Action: What, When, and Where. *Topics in Cognitive Science*, *1*(2), 353–367. https://doi.org/10.1111/j.1756-8765.2009.01024.x

Segaert, K., Poulisse, C., Markiewicz, R., Wheeldon, L., Marchment, D., Adler, Z., Howett, D., Chan, D., & Mazaheri, A. (2022). Detecting impaired language processing in patients with mild cognitive impairment using around-the-ear cEEgrid electrodes. *Psychophysiology*, *59*(5), 1–19. https://doi.org/10.1111/psyp.13964

Segaert, Katrien, Lucas, S. J. E., Burley, C. V., Segaert, P., Milner, A. E., Ryan, M., & Wheeldon, L. (2018). Higher physical fitness levels are associated with less language decline in healthy ageing. *Scientific Reports*, *8*(1), 1–10. https://doi.org/10.1038/s41598-018-24972-1

Segaert, Katrien, Mazaheri, A., & Hagoort, P. (2018). Binding language: structuring sentences through precisely timed oscillatory mechanisms. *European Journal of Neuroscience*, *48*(7), 2651–2662. https://doi.org/10.1111/ejn.13816

Shafto, M. A., & Tyler, L. K. (2014). Decline and Preservation. *Science*, *346*(6209), 583–588. https://doi.org/10.1126/science.1254404

Shahin, A. J., Picton, T. W., & Miller, L. M. (2009). Brain oscillations during semantic evaluation of speech. *Brain and Cognition*, *70*(3), 259–266. https://doi.org/10.1016/j.bandc.2009.02.008

Shaw, D. J., Czekóová, K., Staněk, R., Mareček, R., Urbánek, T., Špalek, J., Kopečková, L., Řezáč, 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*(10896), 1–13. https://doi.org/10.1038/s41598-018-29233-9

Siegel, M., Donner, T. H., & Engel, A. K. (2012). Spectral fingerprints of large-scale neuronal interactions. *Nature Reviews Neuroscience*, *13*(2), 121–134. https://doi.org/10.1038/nrn3137

Smith, T., Gildeh, N., & Holmes, C. (2007). The Montreal cognitive assessment: Validity and utility in a memory clinic setting. *Canadian Journal of Psychiatry*, *52*(5), 329–332. https://doi.org/10.1177/070674370705200508

Stine-Morrow, E. A. L., Soederberg Miller, L. M., & Nevin, J. A. (1999). The effects of context and feedback on age differences in spoken word recognition. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences*, *54*(2), 125–134. https://doi.org/10.1093/geronb/54B.2.P125

Stolk, A., Noordzij, M. L., Verhagen, L., Volman, I., Schoffelen, J. M., Oostenveld, R., Hagoort, P., & 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

Stolk, A., Verhagen, L., Schoffelen, J. M., Oostenveld, R., Blokpoel, M., Hagoort, P., Van Rooij, I., & Toni, I. (2013). Neural mechanisms of communicative innovation. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(36), 14574–14579. https://doi.org/10.1073/pnas.1303170110

Stolk, A., Verhagen, L., & Toni, I. (2016). Conceptual Alignment: How Brains Achieve Mutual Understanding. *Trends in Cognitive Sciences*, *20*(3), 180–191. https://doi.org/10.1016/j.tics.2015.11.007

Takagishi, H., Kameshima, S., Schug, J., Koizumi, M., & Yamagishi, T. (2010). Theory of mind enhances preference for fairness. *Journal of Experimental Child Psychology*, *105*(1–2), 130–137. https://doi.org/10.1016/j.jecp.2009.09.005

Talmi, D., Atkinson, R., & El-Deredy, W. (2013). The feedback-related negativity signals salience prediction errors, not reward prediction errors. *Journal of Neuroscience*, *33*(19), 8264–8269. https://doi.org/10.1523/JNEUROSCI.5695-12.2013

Taylor, J. S. H., Rastle, K., & Davis, M. H. (2013). Can cognitive models explain brain activation during word and pseudoverb reading? A meta-analysis of 36 neuroimaging studies. *Psychological Bulletin*, *139*(4), 766–791. https://doi.org/10.1037/a0030266

Thielmann, I., Hilbig, B. E., & Niedtfeld, I. (2014). Willing to give but not to forgive: Borderline personality features and cooperative behavior. *Journal of Personality Disorders*, *28*(6), 778–795. https://doi.org/10.1521/pedi_2014_28_135

Tognoli, E., Lagarde, J., DeGuzman, G. C., & Kelso, J. A. S. (2007). The phi complex as a neuromarker of human social coordination. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(19), 8190–8195. https://doi.org/10.1073/pnas.0611453104

Tun, P. A. (1998). Fast noisy speech: Age differences in processing rapid speech with background noise. *Psychology and Aging*, *13*(3), 424–434. https://doi.org/10.1037/0882-7974.13.3.424

Tyler, L. K., & Marslen-Wilson, W. (2008). Fronto-temporal brain systems supporting spoken language comprehension. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1493), 1037–1054. https://doi.org/10.1098/rstb.2007.2158

Tyler, L. K., Shafto, M. A., Randall, B., Wright, P., Marslen-Wilson, W. D., & Stamatakis, E. A. (2010). Preserving syntactic processing across the adult life span: The modulation of the frontotemporal language system in the context of age-related atrophy. *Cerebral Cortex*, *20*(2), 352–364. https://doi.org/10.1093/cercor/bhp105

van de Vijver, I., Richard Ridderinkhof, K., & Cohen, M. X. (2011). Frontal oscillatory dynamics predict feedback learning and action adjustment. *Journal of Cognitive Neuroscience*, *23*(12), 4106–4121. https://doi.org/10.1162/jocn_a_00110

van Diepen, R. M., Cohen, M. X., Denys, D., & Mazaheri, A. (2015). Attention and Temporal

Expectations Modulate Power, Not Phase, of Ongoing Alpha Oscillations. *Journal of Cognitive Neuroscience*, *27*(8), 1573–1586. https://doi.org/10.1162/jocn_a_00803

Van Schie, H. T., Mars, R. B., Coles, M. G. H., & Bekkering, H. (2004). Modulation of activity in medial frontal and motor cortices during error observation. *Nature Neuroscience*, *7*(5), 549–554. https://doi.org/10.1038/nn1239

Viana, K. M. P., Zambrana, I. M., Karevold, E. B., & Pons, F. (2016). Beyond conceptual knowledge: The impact of children's Theory-of-Mind on dyadic spatial tasks. *Frontiers in Psychology*, *7*(OCT), 1–11. https://doi.org/10.3389/fpsyg.2016.01635

von Stein, A., Rappelsberger, P., & Sarnthein, J. (1999). Synchronization between temporal and parietal cortex during multimodal object processing in man. *Cerebral Cortex*, *9*, 137–150.

Wang, L., Hagoort, P., & Jensen, O. (2018). Language prediction is reflected by coupling between frontal gamma and posterior alpha oscillations. *Jounral of Cognitive Neuroscience*, *30*(3), 432–447. https://doi.org/10.1162/jocn

Wang, L., Jensen, O., Van den Brink, D., Weder, N., Schoffelen, J. M., Magyari, L., Hagoort, P., & Bastiaansen, M. (2012). Beta oscillations relate to the N400m during language comprehension. *Human Brain Mapping*, *33*(12), 2898–2912. https://doi.org/10.1002/hbm.21410

Wang, L., Zhu, Z., & Bastiaansen, M. (2012). Integration or predictability? A further specification of the functional role of gamma oscillations in language comprehension. *Frontiers in Psychology*, *3*(187), 1–12. https://doi.org/10.3389/fpsyg.2012.00187

Warnell, R. K., & Redcay, E. (2019). Minimal coherence among varied theory of mind measures in childhood and adulthood. *Cognition*, *191*(103997), 1–10. https://doi.org/10.1016/j.cognition.2019.06.009

Waters, G. S., & Caplan, D. (2003). The reliability and stability of verbal working memory measures. *Behavior Research Methods, Instruments, and Computers*, *35*(4), 550–564. https://doi.org/10.3758/BF03195534

Waters, G. S., & Caplan, D. (2005). The relationship between age, processing speed, working memory capacity, and language comprehension. *Memory*, *13*(3–4), 403–413. https://doi.org/10.1080/09658210344000459

Weiss, S., & Mueller, H. M. (2003). The contribution of EEG coherence to the investigation of language. *Brain and Language*, *85*, 325–343. https://doi.org/10.1016/S0093-934X(03)00067-1

Weiss, S., & Mueller, H. M. (2012). "Too many betas do not spoil the broth": The role of beta brain oscillations in language processing. *Frontiers in Psychology*, *3*(JUN), 1–15. https://doi.org/10.3389/fpsyg.2012.00201

Weiss, S., Mueller, H. M., Schack, B., King, J. W., Kutas, M., & Rappelsberger, P. (2005). Increased neuronal communication accompanying sentence comprehension. *International Journal of Psychophysiology*, *57*(2), 129–141. https://doi.org/10.1016/j.ijpsycho.2005.03.013

Weschler, D. (2008). *WAIS-IV administration and scoring manual.* TX: Psychological Corporation.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*(1983), 103–128. https://doi.org/10.1002/app.25226

Wingfield, A., & Grossman, M. (2006). Language and the aging brain: Patterns of neural compensation revealed by functional brain imaging. *Journal of Neurophysiology*, *96*(6), 2830–2839. https://doi.org/10.1152/jn.00628.2006

Wingfield, A., Peelle, J. E., & Grossman, M. (2003). Speech Rate and Syntactic Complexity as Multiplicative Factors in Speech Comprehension by Young and Older Adults. *Aging, Neuropsychology, and Cognition*, *10*(4), 310–322.

https://doi.org/10.1076/anec.10.4.310.28974

Winkler, I., Debener, S., Muller, K. R., & Tangermann, M. (2015). On the influence of high-pass filtering on ICA-based artifact reduction in EEG-ERP. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 2015-Novem*, 4101–4105. https://doi.org/10.1109/EMBC.2015.7319296

Wlotko, E., & Federmeier, K. D. (2012). Age-related changes in the impact of contextual strength on multiple aspects of sentence comprehension. *Psychophysiology*, *49*(6), 770–785. https://doi.org/10.1111/j.1469-8986.2012.01366.x

Wlotko, E., Federmeier, K. D., & Kutas, M. (2012). To predict or not to predict: Age-related differences in the use of sentential context. *Psychology and Aging*, *27*(4), 975. https://doi.org/10.1038/jid.2014.371

Xygalatas, D. (2013). Effects of religious setting on cooperative behavior: A case study from Mauritius. *Religion, Brain and Behavior*, *3*(2), 91–102. https://doi.org/10.1080/2153599X.2012.724547

Zaccarella, E., & Friederici, A. D. (2015). Merge in the human brain: A sub-region based functional investigation in the left pars opercularis. *Frontiers in Psychology*, *6*(NOV), 1–9. https://doi.org/10.3389/fpsyg.2015.01818

Zaccarella, E., Meyer, L., Makuuchi, M., & Friederici, A. D. (2017). Building by Syntax: The Neural Basis of Minimal Linguistic Structures. *Cerebral Cortex*, *27*(1), 411–421. https://doi.org/10.1093/cercor/bhv234

APPENDICES

6.1 Appendix 1: Chapter 2 Supplementary materials

6.1.2 TFR of power between conditions (lexical vs no lexical retrieval & semantic binding vs no semantic binding) for each group (healthy older adults and young adults) separately

Within each age group separately, non-parametric cluster based permutation tests were carried out to compare the time-frequency power changes between the experimental conditions. The experimental manipulations involved: lexical retrieval (real words) vs no lexical retrieval (letter strings); semantic binding (surrounding second word in real word-phrases, including plausible (e.g. swift horse) and implausible (e.g. barking horse) word-phrases) vs no semantic binding (surrounding second word in letter string - real word phrases (e.g. swrfew horse)). The analyses were carried out using the 0 to 3.2sec time window (where 0 is the onset of word 1 and 1.8sec is the onset of word 2), and the pre-defined frequency bands theta (4-7Hz), alpha (8-14Hz), low beta (15-20Hz), and high beta (20-25Hz). We refer to the effects that occurred following word one onset but prior to word two onset as ‘lexical retrieval’ effects. The effects that arose during or post word two onset are referred to as ‘semantic binding’ effects. The supplementary Figure SM.1 displays the condition effects within older healthy adults (A) and young adults (B).

Lexical retrieval within healthy older adults

There was a significantly smaller theta increase in the lexical retrieval condition compared to the non-lexical retrieval condition (p = .02) with a corresponding cluster spanning from 0.3 to 0.65 sec post word one onset, maximal over right occipital and left central electrodes. Secondly, there was a significantly greater theta increase in the lexical retrieval condition compared to the non-lexical condition (p = .004). The observed cluster spanned from 0.8 to 1.35 sec and was maximal over bilateral occipital and central channels.

Furthermore, a cluster in the observed data was found in the alpha band. The lexical condition elicited significantly greater alpha suppression compared to the no lexical retrieval condition (*p* = .012). The observed cluster began around the word 1 onset until 0.55 sec and was maximal over the bilateral parietal-central electrodes. An additional cluster in the alpha range extended from 0.6 to 1.75 sec. The cluster-based permutation testing indicated this cluster to show a significantly greater alpha increase over the occipital and central electrodes in the lexical retrieval condition vs no lexical retrieval (*p* < .001).

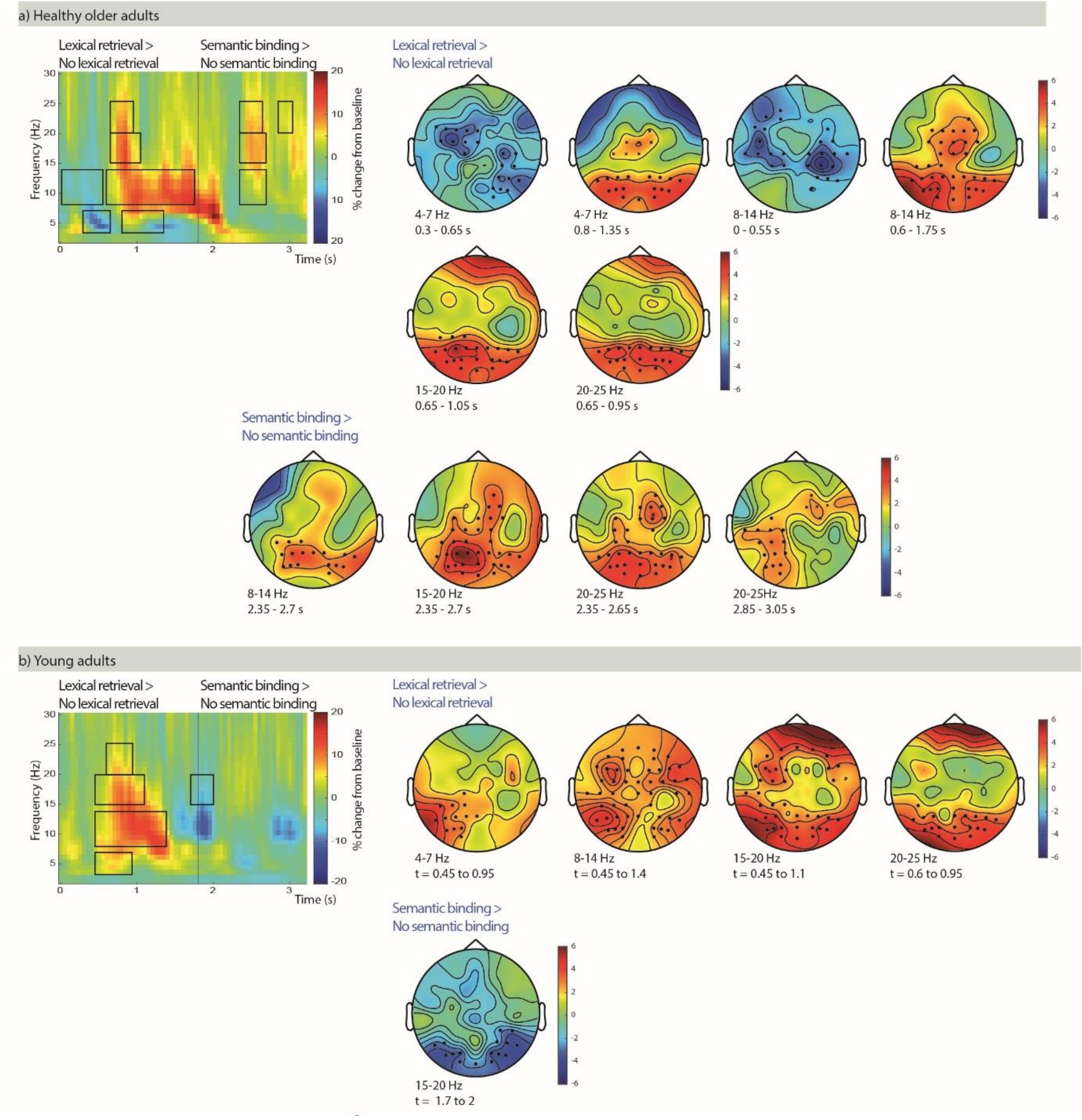


Figure SM.1: TFRs of power for lexical retrieval/semantic binding minus no lexical retrieval/no binding, in (A) the healthy older adults, and (B) the young adults. Head plots are illustrating the clusters of electrodes that show the most pronounced mean condition difference within each group. Black rectangles indicate significant condition differences (*p*< .05, cluster corrected).

Furthermore, a cluster was found in the low-beta band. The observed cluster extended from 0.6 to 1.05 sec and was most pronounced over the occipital electrodes. The cluster-based permutation tests indicated a significant condition effect (*p* = .002) (i.e. significantly less low-beta suppression in the lexical retrieval condition post word one onset compared to the no lexical retrieval condition).

Similarly, we observed a cluster in the high-beta band that extended from 0.65 to 0.95 sec over the occipital channels. The cluster-based permutation test indicated significantly less high-beta suppression in the lexical retrieval compared to the no lexical retrieval condition (*p* = .006).

Lexical retrieval within young adults

The cluster-based permutation tests indicated that there were several significant effects of condition. A cluster in the observed data was found in the theta band that extended from 0.45 to 0.95 sec. There was a significantly greater theta power increase (*p* = .02), maximal over left temporal and parietal channels in the lexical retrieval condition compared to the no lexical retrieval condition.

We also observed a cluster in the alpha band that spanned from 0.45 to 1.45 sec. The cluster corresponded to a significantly smaller alpha power suppression (*p* = .004) in the lexical retrieval compared to the no lexical retrieval condition, which was evident across the whole scalp.

Lastly, clusters in the low-beta and high-beta bands were observed. They extended from 0.45 to 1.1 sec maximal over frontal, occipital and left temporal electrodes and from 0.6 to 0.95 sec, maximal over occipital electrodes respectively. The clusters corresponded to significant condition effects indicating less low-beta and high-beta suppression in the lexical retrieval condition compared to the no lexical retrieval condition (*p* = .002 and *p* = .002 respectively).

Semantic binding within healthy older adults

The cluster-based permutation tests found several significant condition effects. A cluster in the observed data was found in the alpha band that extended from 2.35 to 2.7 sec over occipital and parietal electrodes. There was a significantly reduced alpha decrease (*p* = .038) in the semantic binding condition compared to the no semantic binding condition.

Moreover, clusters in the observed data were found in the low-beta and high-beta ranges spanning from 2.35 sec and ending around 2.65-2.7 sec over occipital, parietal and central electrodes. There was a significantly smaller low-beta (*p* = .002) and high-beta (*p* =.002) in the semantic binding condition compared to the no semantic binding condition.

Further condition effects corresponded to a cluster extending from 2.85 to 3.05 sec in the high-beta (maximal over left occipito-central channels) bands. There was a significantly greater high-beta increase in the semantic binding condition compared to no semantic binding (*p* =.044).

Semantic binding within young adults

The cluster-based permutation tests indicated a significant condition effect. The observed cluster was found in the low-beta band extending around the second word onset (1.7 to 2 sec) and was maximal over the occipital channels. The condition effect revealed a significantly smaller low-beta increase (*p* = .036) in the semantic binding condition compared to the no semantic binding condition.

6.1.3 Full report of the sub-sampling results

In order to rule out that the observed condition effects reported in the main paper could be attributed to the difference in the number of trials between conditions, we re-did our analysis matching the number of trials (i.e. random re-sampling) between conditions. We still were able to observe all our previous effects when matching the trial numbers between conditions. We report the results of the re-sampled data below (see also supplementary Figure SM.2).

TFR of power between conditions (lexical vs no lexical retrieval & semantic binding vs no semantic binding) for each group (healthy older adults and young adults) separately with the re-sampled data set

Lexical retrieval within healthy older adults

There was a significantly smaller theta increase in the lexical retrieval condition compared to the non-lexical retrieval condition (*p* < .001) with a corresponding cluster spanning from 0 to 0.6 sec post word one onset, evident across the whole scalp. Secondly, there was a significantly greater theta increase in the lexical retrieval condition compared to the non-lexical condition (*p* = .009). The observed cluster spanned from 0.8 to 1.35 sec and was maximal over bilateral occipital and central channels.

Furthermore, a cluster in the observed data was found in the alpha band. The lexical condition elicited significantly greater alpha suppression compared to the no lexical retrieval condition (*p* = .02). The observed cluster began around the word 1 onset until 0.55 sec and was evident across the whole scalp. An additional cluster in the alpha range extended from 0.6 to 1.7 sec. The cluster-based permutation testing indicated this cluster to show a significantly greater alpha increase over the occipital and central electrodes in the lexical retrieval condition vs no lexical retrieval (*p* < .001).

Furthermore, a cluster was found in the low-beta band. The observed cluster extended from 0.65 to 1.05 sec and was most pronounced over the occipital electrodes. The cluster-based permutation tests indicated a significant condition effect (*p* < .001) (i.e. significantly less low-beta suppression in the lexical retrieval condition post word one onset compared to the no lexical retrieval condition).

Similarly, we observed a cluster in the high-beta band that extended from 0.65 to 0.85 sec over the occipital channels. The cluster-based permutation test indicated significantly less high-beta suppression in the lexical retrieval compared to the no lexical retrieval condition (*p* = .002).

Lexical retrieval within young adults

The cluster-based permutation tests indicated that there were several significant effects of condition. A cluster in the observed data was found in the theta band that extended from 0.35 to 0.9 sec. There was a significantly greater theta power increase (*p* = .007), maximal over left temporal and parietal channels in the lexical retrieval condition compared to the no lexical retrieval condition.

We also observed a cluster in the alpha band that spanned from 0.45 to 1.4 sec. The cluster corresponded to a significantly smaller alpha power suppression (*p* < .001) in the lexical retrieval compared to the no lexical retrieval condition, which was evident across the whole scalp.

Lastly, clusters in the low-beta and high-beta bands were observed. They extended from 0.5 to 1.15 sec maximal over frontal, occipital and left temporal electrodes and from 0.5 to 1.15 sec, maximal over occipital electrodes respectively. The clusters corresponded to significant condition effects indicating less low-beta and high-beta suppression in the lexical retrieval condition compared to the no lexical retrieval condition (*p* < .001 for both effects).

Semantic binding within healthy older adults

The cluster-based permutation tests found several significant condition effects. Firstly, there was a significantly smaller theta increase in the semantic binding condition compared to the no binding condition (*p* = .003) with a corresponding cluster spanning from 1.95 to 2.4 sec, evident across the whole scalp. Furthermore, a cluster in the observed data was found in the alpha band that extended from 2.35 to 2.85 sec over occipital and parietal electrodes. There was a significantly reduced alpha decrease (*p* = .011) in the semantic binding condition compared to the no semantic binding condition.

Moreover, clusters in the observed data were found in the low-beta and high-beta ranges spanning from 2.35 sec and ending around 2.6-2.75 sec over occipital, parietal and central electrodes. There was a significantly smaller low-beta (*p* < .001) and high-beta (*p* =.002) in the semantic binding condition compared to the no semantic binding condition.

Further condition effects corresponded to a cluster extending from 2.85 to 3.05 sec in the high-beta (maximal over left occipito-central channels) bands. There was a significantly greater high-beta increase in the semantic binding condition compared to no semantic binding (*p* =.013).

Semantic binding within young adults

The cluster-based permutation tests indicated significant condition effects. The observed clusters were found in the alpha, the low- and high-beta bands extending around the second word onset (1.7 to 2 sec) and were maximal over the occipital channels. The condition effects revealed significantly smaller alpha (*p* = .048), low-beta (*p* = .018), and high-beta (*p* = .017) increases in the semantic binding condition compared to the no semantic binding condition.

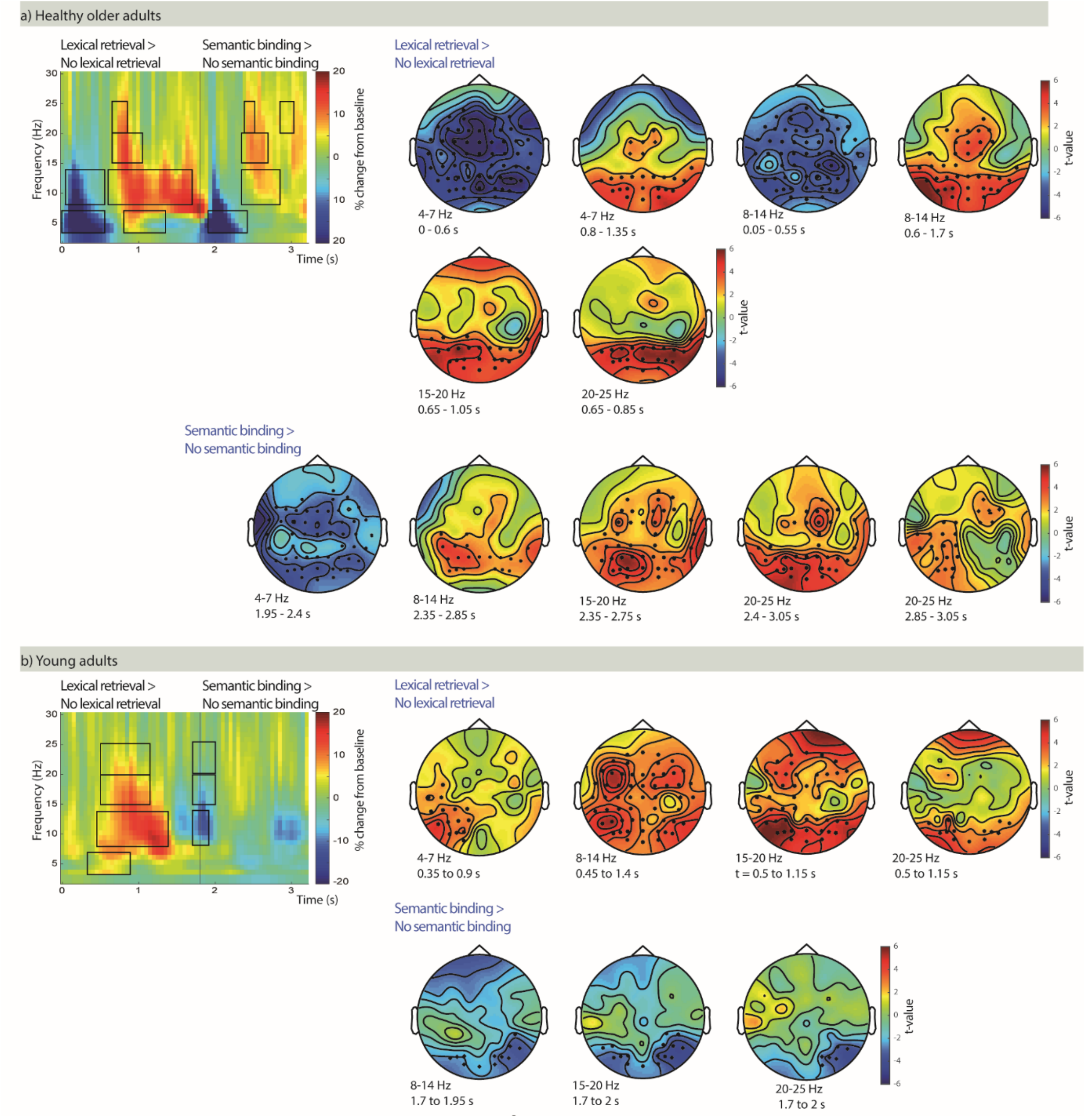


Figure SM.2: TFRs of power for lexical retrieval/semantic binding minus no lexical retrieval/no binding with re-sampled data, in (A) the healthy older adults, and (B) the young adults. Head plots are illustrating the clusters of electrodes that show the most pronounced mean condition difference within each group. Black rectangles indicate significant condition differences (*p*<0.05, cluster corrected).

6.2 Appendix 2: Chapter 3 Supplementary material

6.2.1 Interplay of behavioural adjustment (between players) was predictive of cooperative outcome

We sought to examine whether the outcome of each trial could be predicted based on the interplay of adjustments made by both players on a behavioural level. However, the nature of our experimental design, which involved random interleaving of condition types within blocks rather than a block-based structure, prevented us from directly examining sequential behavioural adjustments between consecutive trials. This constraint arose because the specific condition type consistently influenced the 'change' or 'adjustment' in reaction time (RT) due to the varying wait times required (short, medium, or long).

To overcome this, we established a reference point, or 'baseline,' to facilitate the calculation of behavioural adjustments for each trial. Our reasoning centred on the notion that 'learning' occurred immediately after the present trial, coinciding with participants receiving feedback. Consequently, for each subsequent trial following this feedback, we identified the corresponding condition type (signifying the moment of adjustment). We then located the nearest preceding trial with the same condition type, which served as our baseline. Using this approach, we computed the behavioural adjustment for each player. This adjustment was quantified as the difference between the RT of the previous trial with the same condition type and the RT of the subsequent trial with the same condition type.

With this, we were then able to fit a multiple linear regression model (estimated using Ordinary Least Squares), with the predictors being: the behavioural adjustment of player one and behavioural adjustment of player two. The outcome variable was the cooperative outcome (i.e., the time difference between the button press of player one and two); cooperative outcome ~ behavioural adjustment of player one * behavioural adjustment of player two. This was carried out on a trial-by-trial basis.

The model explained a statistically significant but weak proportion of variance (F(3, 2647) = 61.40, $p < .001$).The model revealed that the behavioural adjustment of player one ($\beta = -0.14$, $p < .001$) but not of player two ($\beta = -1.11$, $p = 0.945$) was a significant predictor of cooperative outcome. Most importantly, the interaction between behavioural adjustment of player one and behavioural adjustment of player two significantly predicted cooperative outcome ($\beta = -0.16$, $p < .001$).

While this model indeed confirms the significance of the interplay between the behavioural adjustments made by both players as a strong predictor of continuous cooperative outcomes (i.e., the time gap between player one's and player two's button presses), it is important to exercise caution when interpreting these findings. Our experimental design posed limitations regarding the direct calculation of behavioural adjustments occurring from one trial to the next due to the presence of intermixed trial types, making trial type (short, medium, and long) a confounding variable. Consequently, we had to establish a baseline for behavioural adjustment using the previous trial of the same condition and calculate *behavioural adjustment*, which may not fully represent the way that participants updated their mental representations of the timing intervals on the next consecutive trial. Therefore, it is important that these results are replicated using a blocked experimental design, enabling to directly examine the effect of behavioural adjustment of both players on cooperative outcome.

6.2.2 Feedback related results

Feedback related oscillatory modulations in theta, alpha and beta bands

We compared the feedback related activity, which includes evoked and induced activity, between correct and incorrect conditions across all of the participants (N =59) aligned to the onset of the feedback text. Suppl. Figure S.M.3A displays the TFRs of power locked to the onset of the feedback text for the incorrect, and correct conditions, the difference between them, as well as the topographic distributions of the condition effects. Below we describe the significant oscillatory differences between correct (i.e., successful cooperation) and incorrect (i.e., failed cooperation) conditions that were identified using non-parametric cluster-based permutation tests using a 0-4sec time window post feedback text onset. We observed significant condition differences in the pre-defined delta (1-4Hz), theta (4-7Hz), alpha (8-14Hz), low beta (15-20Hz) and high beta (20-25Hz) frequency bands following the onset of the feedback text. Firstly, we observed significant condition differences (p = .003 and p <.001 respectively) in theta and delta activity maximal over the central and centro-parietal channels. Specifically, theta (and delta) power was stronger right after the onset of the feedback text in the incorrect condition until around 0.4sec compared to the correct condition. Further, there was a significant condition difference (p < .001), with greater delta power in the incorrect vs. correct condition. This effect corresponded to a cluster maximal over the central electrodes that stretched between 1.15 to 1.75sec post feedback text onset. We consider this effect to be a representation of the sustained ERP in the similar time window. Finally for the delta band, we found significantly more pronounced activity in the incorrect compared to the correct condition following the feedback onset (p = .006). This effect was most prominent between 2.65sec and 2.75sec post feedback text onset and maximal over right centro-temporal channels. Further, in the theta band, there was a more sustained theta attenuation in the incorrect compared to the correct condition (p < .001). This effect was maximal over centro-parietal and parietal sites and most prominent between 0.6 to 1.15sec post feedback text onset. Lastly for the theta band, we observed a condition effect that was maximal over temporal electrodes and left lateralised (p = .039). Specifically, there was a reduction in theta power in the incorrect condition relative to the correct condition corresponding to a cluster that stretched between 1.5 and 1.6sec post feedback onset. We also found significant condition differences in alpha activity. Firstly, there was a greater alpha power decrease in the incorrect condition relative to the correct condition (p < .001). This effect corresponded to a cluster that extended between 0.3 to 1.75sec post feedback and was maximal over central electrodes. We also found that there was a greater alpha power increase in the correct condition relative to the incorrect condition (p = .031). This effect corresponded to a cluster that spread between 2 to 2.25sec and was maximal over left temporal channels.

We found a significant condition effect in the low beta band (p =.024). This corresponded to a transient time window of .2 to .25sec post feedback and was maximal over centro-parietal electrodes. Further, there was a prolonged opposite pattern of low beta (p < .001) and high beta (p < .001) activity between the conditions. The incorrect condition yielded a prolonged low beta decrease, whereas the correct condition showed a prolonged low beta increase. This effect corresponded to a cluster that spanned from .35 to 1.5sec and was prominent all over the scalp. Lastly, we found a transient condition effect (p = .013) in the high beta band that extended between 2.9 and 3sec post feedback. Specifically, there was a greater high beta power increase in the incorrect compared to the correct condition, maximal over bilateral temporal sites.

Initial sustained ERP positivity and later negativity associated with incorrect trials after feedback

We compared the ERPs between incorrect and correct conditions across all of the participants (N =59) using a time window of 0 to 4sec post feedback text onset. Suppl. Figure S.M.3B shows feedback locked averaged ERP waveforms for incorrect and correct conditions, as well as the condition effect topographic distributions. The cluster-based permutation tests revealed that there was a significant effect of condition, specifically the incorrect condition was associated with a greater positivity compared to the correct condition (p < .001). This effect corresponded to a cluster that spanned from .178 to .926sec and was maximal over centro-parietal electrodes. Secondly, we found another condition effect, where the incorrect trials yielded a sustained negativity ERP relative to the correct trials (p < .001, p

= .034). This corresponded to clusters that extended from 1.406 to 2.554sec as well as 2.562 to 2.634sec respectively, most prominent over centro-parietal channels.

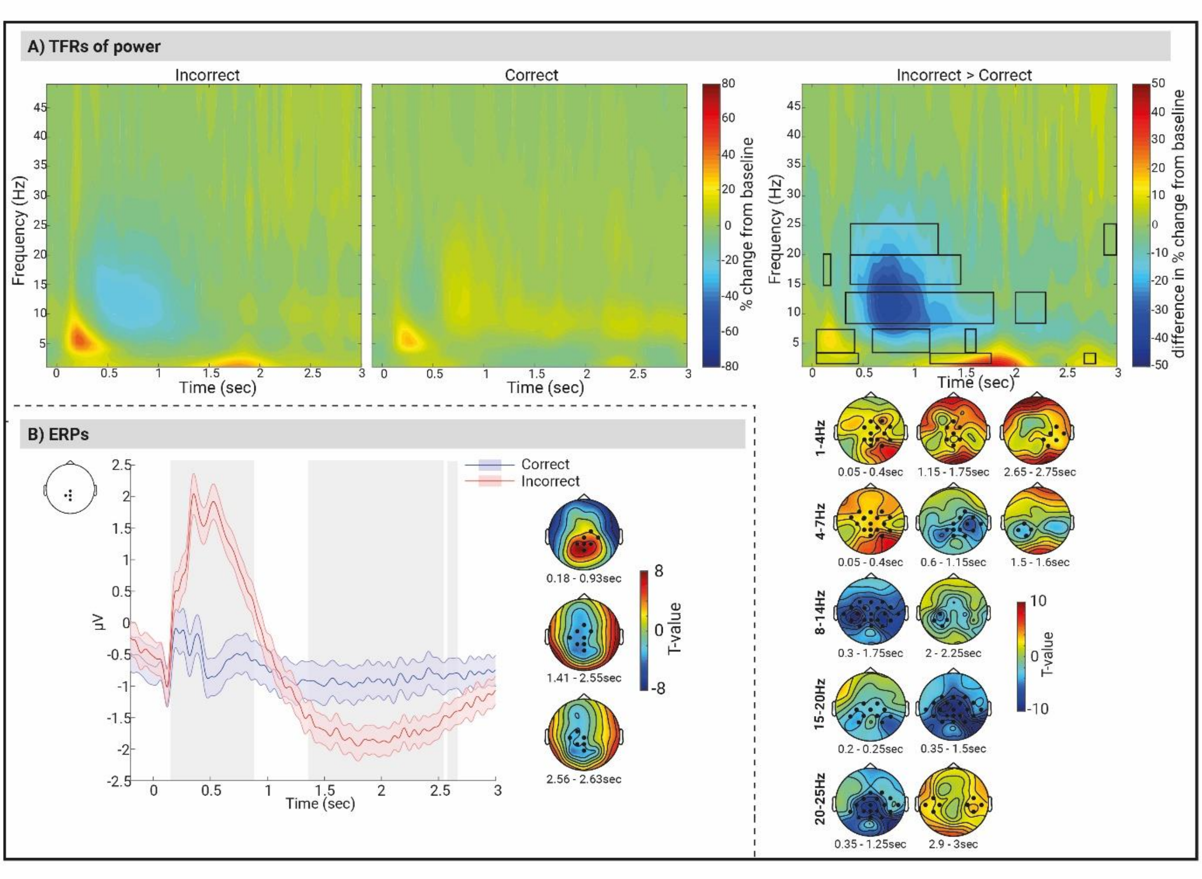


Figure SM.3: (A) TFRs of power (at the Cz electrode) for incorrect trials (left), correct trials (middle), and the condition differences (i.e., incorrect – correct) (right) averaged over all individuals (N = 59), locked to the onset of feedback text. Headplots illustrate the clusters of electrodes that show the most pronounced mean condition difference. Black rectangles indicate significant condition differences ($p < 0.05$, cluster corrected). (B) Feedback locked averaged ERPs produced by correct (blue) and incorrect (red) trials in the timing task averaged over all individuals (N = 59). The ERP waveforms show averaged ERPs across the overlapping (across the three time windows of the significant between condition effects) electrode clusters (Cz, CP1, Pz, and CPz) that indicate the maximal condition difference (a schematic view of these electrodes is shown in the top right corner). The shaded areas around the ERP waves represent standard error. The grey rectangles represent time windows of the significant between condition differences. The black dots in the headplots illustrate the clusters of electrodes showing most pronounced mean condition differences for each time window.

6.2.3 Effects of brain-to-brain coupling on global cooperative outcomes analysis

We performed further analysis to identify power brain-to-brain signatures between participant pairs that were predictive of *global* cooperative outcomes. In our experimental design, condition types (i.e., high/short, medium, and low/long) were interleaved in a randomised manner within each experimental block, meaning that successive trials frequently did not share the same condition type. Given this setup, we aimed to explore the potential for predicting cooperative outcomes of subsequent trials with the same condition types based on inter-player neural signatures – we refer to as the "global effect".

Our approach focused on examining the nearest trial of the same condition type, instead of studying consecutive trials. Specifically, we categorized unsuccessful trials into two distinct conditions: forecasting unsuccessful cooperation on the nearest trial of the same condition type (involving failed cooperation trials followed by another unsuccessful cooperation trial of the same condition type), and forecasting successful cooperation on the nearest trial of the same condition type (entailing failed cooperation trials followed by a successful cooperation trial in the subsequent non-consecutive instance). The average number of trials included in the analysis was 41.28 (SD = 12.27) for the condition of forecasting successful cooperation and 53.45 (SD = 29.58) for the condition of forecasting failed cooperation. In order to compare the two conditions we used non-parametric cluster-based permutation tests across all pairs (N = 29 pairs). We followed the exact same steps as described in the main manuscript section "*Effects of brain-to-brain coupling on local cooperative outcomes analysis*".

Inter-player power dynamics did not forecast cooperative results in the following trial of the same condition

No significant inter-player power coupling patterns were observed when forecasting cooperative outcomes for the nearest trial of the same condition type within the theta band (utilizing the predefined time window of 0-0.4 seconds), nor within the alpha band (utilizing the predetermined time window of 0.3 to 1.75 seconds). Subsequent exploratory analyses involving a time window of 0 to 2 seconds post feedback similarly failed to reveal significant effects across any of the examined frequency bands, including alpha, low beta, and high beta.

6.3 Appendix 3: Chapter 4 Supplementary Material

6.3.1 Communicative cooperation time ran out scores and individual Theory of Mind

A forward multiple regression analysis was carried out to identify possible predictors of time ran out scores out of the following candidate variables: ToM score of player one, ToM score of player two, Raven score of player one, Raven score of player two, and ToM score difference between player 1 and player 2 within each pair. The regression model revealed that ToM competence (i.e. ToM score of player one or ToM score of player two) was not a significant predictor for time ran out scores.

Table SM.1:. Correlation matrix for outcome (cooperative failure, cooperative success, and time ran out) and predictor variables in the regression analysis. Values in brackets represent the 95% confidence lower and upper intervals.

	Failed cooperation	Successful cooperation	Time ran out	Player 1 ToM score	Player 1 Raven score	Player 2 ToM score	Player 2 Raven score
Player 1 ToM score	-.177* (-.32, -.03)	.098 (-.05, .24)	-.015 (-.16, .13)	-			
Player 1 Raven's score	-.139* (-.28, .01)	.078 (-.07, .22)	-.013 (-.16, .14)	.274*** (.13, .41)	-		
Player 2 ToM score	-.162* (-.30, -.01)	.130* (-.02, .27)	-.051 (-.20, .10)	.060 (-.09, .21)	.104 (-.04, .25)	-	
Player 2 Raven's score	-.045 (-.19, .10)	.180** (.03, .32)	-.146* (-.29, .01)	.102 (-.05, .25)	0.26 (-.12, .17)	.267*** (.12, .40)	-

p*<.05, ** *p*<.01, * *p*<.001

Here, we also report the results from the median split analysis approach, in which we split our sample into low and high ToM groups based on individual MASC scores. A one-way between participants ANCOVA was conducted to assess for quantitative differences between the ToM groups (i.e., low ToM and high ToM) on the time ran out scores whilst adjusting for general ability (i.e., Raven score). The ANCOVA with time ran out score as the dependent variable, using the Raven's score as the covariate, did not reveal a significant main effect of ToM group upon time ran out scores, $F(1, 347) = .053$, $p = .818$, $\eta^2_p = .056$. After adjusting for Raven score, the time ran out adjusted mean % for the low ToM group was 41.415 (SEM = 1.094), and for the high ToM group was 41.33 (SEM = 1.216) (see Figure S.M.4). This is to be expected as the time ran out scores contain a mixture of 'almost' successful and 'almost' failed cooperation.

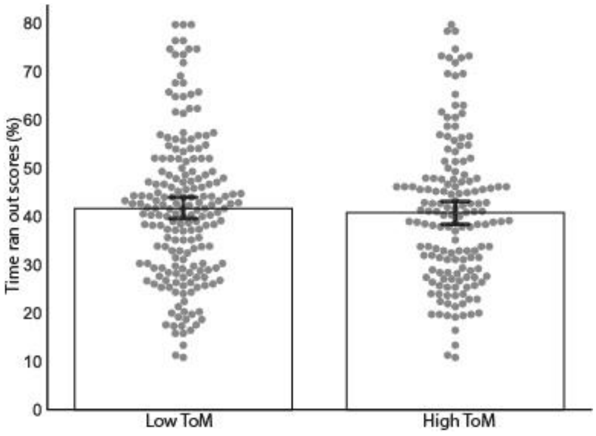


Figure SM.4: Bar graphs for mean % score for cooperative time ran out scores showing each group performance separately (low and high Theory of Mind groups). Error bars represent 95% confidence intervals. *Please note that the outcome variable (i.e., time ran out scores) is duplicated here (if one participant was in the high ToM group and their partner was in the low ToM group, their dyadic time ran out score is plotted twice in the figure, one on each x-axis bar).*

6.3.2 Communicative time ran out scores and Theory of Mind of both partners within a pair

A one-way between participants ANOVA did not yield any significant differences between groups on the time ran out scores ($F(2,172) = .578$, $p = .562$). Again, this is to be expected as the time ran out responses contain pure noise (i.e., a mixture of ‘almost’ failed and ‘almost’ successful cooperation) (see Figure S.M.5).

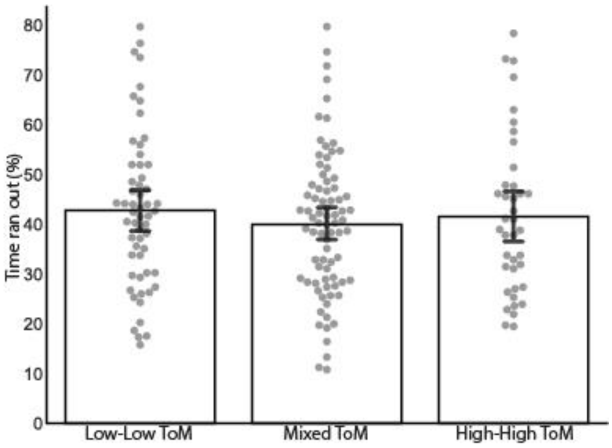


Figure SM.5: Bar graphs for mean % score for cooperative success (panel A) and cooperative time ran out (panel B) showing each group performance separately (low-low, high-high and mixed Theory of Mind groups). Error bars represent 95% confidence intervals.