Kent Academic Repository

De Lillo, Martina (2021) ***Social cognition across the lifespan and its relation to executive functions.*** **Doctor of Philosophy (PhD) thesis, University of Kent,.**

SOCIAL COGNITION ACROSS THE LIFESPAN AND ITS RELATION TO EXECUTIVE FUNCTIONS

Martina De Lillo

A thesis submitted for the degree of Ph.D. in the Faculty of Social Sciences at the

University of Kent, *September 2020*

Declaration

The research presented in this thesis was conducted at the School of Psychology, University of Kent, whilst the author was a full-time postgraduate student, supported by a PhD studentship from the European Research Council grant (Ref: CogSoCoAGE; 636458), under the supervision of Professor Heather Ferguson. Some data from Chapter 2 and all of Chapter 3 of this thesis are under review at different journals (please see below). The data reported in Chapters 2, 3, and 4 has been presented at numerous conferences.

Presentations and forthcoming publications

Papers under review:

- De Lillo, M., Brunsdon, V., Bradford, E., Gasking, F., & Ferguson, H. J. (under review). Training executive functions using a 21-day adaptive procedure and active control group. *Quarterly Journal of Experimental Psychology*.
- De Lillo, M., Foley, R., Fysh, M., Stimson, A., Bradford, E., Woodrow-Hill, C., & Ferguson, H. J. (under review). Tracking developmental differences in real-world social attention across adolescence, young adulthood and older adulthood. *Nature Human Behaviour*.

Book chapter:

- Bradford, E., De Lillo, M., & Ferguson, H. J. (in press). Chapter 9. The future of research on social interaction. In Ferguson, H.J. & Bradford, E. (Eds.), *The cognitive basis of social interaction across the lifespan*, Oxford University Press.

Papers in preparation:

- De Lillo, M., Woodrow-Hill, C., Foley, R., Bradford, E., & Ferguson, H. J. Empathy for others' physical and social pain in adolescence, young adulthood and older adulthood.

Invited talk:

- De Lillo, M. Training cognitivo: dalle funzioni esecutive alla cognizione sociale. Presented at *Universita' di Bari, May 2019.*

Peer-reviewed presentations (talks and posters):

- De Lillo, M., Brunsdon, V.E.A., Bradford, E.E.F., Fysh M., & Ferguson, H. Investigating social attention across the lifespan, Presented at *Experimental Psychology Society, January 2020, London.*
- De Lillo, M., Brunsdon, V.E.A., Bradford, E.E.F., Fysh, M., & Ferguson, H. Investigating attention to social stimuli in real world in Younger *vs.* Older adults. Presented at *"From Self-knowledge to knowing others" workshop, Brussels Thursday 7th of November 2019, UCLouvain.*
- De Lillo, M., Brunsdon, V.E.A., & Bradford, E.E.F & Ferguson, H. Comparing EEG activity in response to empathy for social vs physical pain. Presented at *Experimental Psychology Society, July 2019, Bournemouth.*
- De Lillo, M., Brunsdon, V.E.A., Bradford, E.E.F., Fysh, M., & Ferguson, H. Comparing real-time social interactions in young and old adults. Presented at *Experimental Psychology Society, April 2019, Manchester.*
- De Lillo, M., Brunsdon, V.E.A., Bradford, E.E.F., Fysh, M., & Ferguson, H. Inferring Mental States in the Lab and in Real Life: Evidence From Younger and Older Adults. Presented at *the Psychonomic Society Annual Meeting, New Orleans, USA.*

- De Lillo, M., Brunsdon, V.E.A., Bradford, E.E.F., Fysh, M., & Ferguson, H. An investigation on Executive Functions and Theory of Mind across the Lifespan. Presented at *ERC Conference – Social Communication Across the Lifespan Canterbury, Kent, UK.*
- De Lillo, M., Brunsdon, V.E.A., Bradford, E.E.F., & Ferguson, H. (2018). An eye-tracking investigation of real-time social interaction. Presented at *the Psychonomic Society Annual Meeting, Amsterdam, The Netherlands.*
- De Lillo, M., Brunsdon, V.E.A., & Bradford, E.E.F & Ferguson, H.J. Training cognitive abilities using a 21-day adaptive procedure. Presented *at Experimental Psychology Society, January 2018, London.*
- Brunsdon, V.E.A., Bradford, E.E.F., De Lillo, M., & Ferguson, H.J. *Changes in executive function across adulthood.* Presented at *the International Conference for Cognitive Neuroscience of Executive Functions, September, 2017, Padova.*

Abstract

Successful social interactions represent a crucial aspect of our everyday life. To achieve this, we have to interpret and understand others’ thoughts and intentions, take their perspective, and empathize with them. All these processes characterize a wider concept called “social cognition”. This thesis aims to shed light on social cognition across the lifespan by adopting a large battery of tasks that included behavioural, EEG, and real-world eye-tracking measures. Starting from the hypothesis that social cognition is impaired in adolescence and old age, I compared the performance of these two age groups with young adults on measures of perspective-taking, empathy and social attention. Results revealed that age modulates different sub-components of social cognition in distinct ways; more cognitively demanding components are more likely to suffer a decline in older age. In addition, due to the link between social cognition and executive functions, this work also tested whether cognitive and social abilities can be enhanced indirectly through cognitive training. In particular, I designed a 21-day cognitive training protocol that targets Working Memory, Inhibitory Control, or Cognitive Flexibility (versus an active control group) to test whether improvements can be detected on the training task (direct transfers) or on a task that measures the same cognitive ability (near transfers), and test the generalisability of these effects to other cognitive and social constructs (far transfers). Whereas robust direct training effects emerged, limited near transfers and no far transfers were detected.

Acknowledgements

Firstly, I’d like to thank my great supervisor Professor Heather Ferguson for giving me the opportunity to pursue this PhD. She supported me at each step of this exciting journey and gave me all the resources to grow professionally. She always found the time when I needed it despite her super busy schedule. She is absolutely a “Wonder Woman”. I would also like to thank Lizzie and Viki for all the professional and personal support they gave me in these years. Thanks also to Aimee, Brittany, Becca, and Millie who have helped me to collect data for this project. Without them it would have been much harder. A big thanks to all the participants who dedicated their time to participate in this project, and to the ERC for funding my PhD and this project.

I want to thank my family, my father Francesco, my mother Rosaria, my sister Marcella, who have always supported me in many ways through my education. I’d like to thank Jon for being that special support in these years and all my amazing friends, Mahsa, Silvia, Matt, Giulia, Courtney, Andre, Ana, and Maria, who made this journey more bearable with laughs and fun nights.

Last but not least, I’d like to thank my ‘far people’, my auntie Silvana, Roberta, Ilaria, Roberto, Sara, and Marina, for the long audios and calls in these years. Because of you I felt closer to home.

Table of Contents

Table of Figures

List of Tables

CHAPTER 1: INTRODUCTION

1.1. Overview

Predicting our friends' behaviour, understanding the body language of our partner, feeling upset after watching an athlete losing an important match, trying to consider our interlocutors' perspective. All these scenarios involve a different set of social skills, such as inferring other's mental states, considering their points of view and sharing their emotions. We continuously engage these processes in our everyday life and they constitute the basis of our social interactions. Interpreting our social world is not always easy, and under certain conditions, can involve a great deal of effort. Social behaviour has a complex nature that not only requires us to understand and to interpret observable behaviour, but also to perceive and to infer mental states and emotions (Frith & Frith, 2007; Kanske & Murray, 2019). In order to understand what is happening around us, we have to deal with a range of interactions that range from verbal communications to more complex types of understanding, such as interpreting facial expressions and body gestures.

Social interactions are not just characterized by what we can perceive from others. In fact, we are also influenced by our emotional status: our interactions are driven by our own emotions and our feelings. How we feel in a particular moment, can influence how we react and what we can perceive when interacting with others. For example, if we are in a bad mood, we can perceive a behaviour in a completely different way than if we were in a good mood. At the same time, our perceptiveness can also be affected by our cognitive functioning (i.e. ability to memorize information, to pay attention on relevant stimuli; Fiske & Taylor, 1991). If we imagine a typical conversation in which we do not pay attention to what our interlocutor has just communicated to us, we might respond incoherently, or even inadvertently offend them. If we are not able to switch between two sources of information, we might confuse their meaning or intentions. If we are unable to inhibit an impulsive behaviour or response, we could hurt ourselves and others. Successfully managing everyday life situations such as these involves underlying cognitive control processes, such as memory, flexibility and inhibitory control, which play key roles in facilitating higher order processes to understand others' mental states (Apperly, 2010).

The processes that characterize social interactions, include two broad domains: one social (e.g. sharing others' emotions) and one cognitive (e.g. inhibiting a response). The central aim of this thesis is to better characterise how social abilities change across the lifespan (comparing adolescents, young and older adults), and to explore the relationship between social and cognitive functioning by testing whether social cognitive abilities can be enhanced indirectly by training the underlying cognitive mechanisms (specifically, working memory, inhibitory control, and cognitive flexibility). In this introductory chapter, I will first summarise the prominent theories of social cognition, with a particular emphasis on the concept of Theory of Mind. Second, I will illustrate the relationship between social abilities and cognitive skills, and discuss empirical evidence that supports this link. Finally, I will explore how these cognitive abilities and social cognition skills develop across the lifespan.

1.2. Theories of Social Cognition

Social cognition has been defined as a series of psychological processes that support social understanding (Adolph, 2009). According to Frith (2008), processes such as perception, attention, memory and action planning "are important in social interactions and the study of information processing in a social setting is referred to as social cognition". Despite the important role of these abilities, currently there is still no agreement on the specific processes that characterize social cognition, or in what ways they are interrelated with each other (Happé, et al., 2017). In fact, as indicated by Happé et al. (2017), there is quite a variance of the number of components to be included in social cognition: Fiske and Taylor (2013) refer to 14 domains, starting from more basic abilities such as social memory and social attention, to more complex concepts like social inferences and decision making; whereas Happé and Frith (2014) identify 10 domains, including empathy, self-processing, affiliation. There is however, a common agreement that social cognition helps us to understand ourselves, others and the environment where we act, through implicit and explicit processes (Moskowitz & Okten, 2017).

Generally, studies on social cognition strive to answer questions such as: how we represent other's *perspective*, whether we can e*mpathize*, and whether we *understand* and *predict* others' actions/ thoughts. Hence, social cognition includes a range of capacities among which are empathy, perspective-taking, making social inferences and mindreading. It is important to mention that in the literature on this field, these terms have often been used interchangeably and they are often referred to under a more general 'umbrella' concept, known as Theory of Mind (ToM). Specifically, ToM is the ability to understand and predict others' mental states, including their desires, beliefs, emotions and knowledge (Premack & Woodruff, 1978; Baron-Cohen, 1997). This term also includes the concept of mentalizing, that is, the understanding we have about people having a mind, and the notion of mindreading, which is the ability to comprehend others' behaviours and predict their actions by putting ourselves 'in their shoes'. ToM processes can be synthetized as three crucial abilities: i) comprehend and understand others' behaviours, beliefs, intentions, ii) predict others' behaviour, and iii) manipulate all this information (Poletti, et al., 2012).

Given its importance, over across the last decades, diverse theories have tried to explain how ToM works. Traditional theories of ToM have leaned heavily on philosophy of mind, defining observers' key task as to reason about the relationship between an agent's intentions, beliefs and desires and their actions. Davies and Stone (1995) describe ToM through a model called "Theory-theories"; the authors suggest that ToM refers to the ability of having a concept of mental state, and having beliefs about others' mental states by observing the world and updating their theory about others' minds through experience. "Theory-theories", however, were not restricted to explaining people and their minds, they also applied to objects. In contrast to "Theory-theories", "Simulation theories" adopt the idea that we possess our own concept of mind and through that idea we simulate the mind of others (Gallese & Goldman, 1998). This theory is supported by the existence of the mirror neuron system (MNS): a series of neurons located in the prefrontal motor cortex, the inferior parietal gyrus and the superior temporal sulcus, that fire when we observe someone's actions as well as when we perform those actions ourselves (Rizzolatti & Craighero, 2004; Rizzolatti, et al., 2009). This theory will be further detailed in Chapter 2, however here it is important to note that the activation of the mirror neuron system has been proposed as a route through which we understand other's behaviours and intentions- by mentally simulating the same patterns of actions. Thus, imitation observation of actions helps to take the perspective of another person (Catmur, et al., 2007).

Later, Leslie et al., (2004) speculated that our mind is formed of two separate mechanisms that collaborate to provide ToM, therefore we are able to formulate multiple possible intentions and beliefs. The theory of mind mechanism (ToMM) allows people to generate and represent multiple possible beliefs. A more recent theory has been developed by Gallagher and Hutto (2008), affirming that when we experience real-world social interactions, we understand others' intentions through their expressive behaviour. Therefore, the authors postulate the existence of an "interaction theory", which is also sustained by the MNS.

Among the first to detail a theoretical framework for ToM that links social abilities with cognitive mechanisms is the 'two systems model for mindreading', proposed by Apperly and Butterfill (2009, 2011; see also Apperly, 2009). According to this account, two systems exist for belief reasoning: one is automatic, inflexible and cognitively efficient (and hence reflects animals' and infants' basic ToM, and adults' moment-by-moment social cognition), and the other is more flexible but is cognitively demanding (and therefore more suited for explicit and planned ToM inferences). This two-systems model predicts that while some aspects of ToM performance will correlate with changes in executive functions and other cognitive abilities (e.g. belief reasoning and inferences from language), those that tap into the cognitively efficient system 1 (e.g. emotion reading and visual perspective-taking) should not reveal a comparable change with reduced cognitive abilities. The link between social and cognitive capacities is further supported by a model proposed by Carruthers (2016), however this model proposes that a single mindreading system can account for the relationship between EFs and mindreading abilities, and that this is shaped through gradual conceptual enrichment so that EFs can be recruited (or not) to support ToM according to need.

To date, the debate about the functioning of ToM remains open, however, there is a large consensus that two domains characterize ToM: a cognitive one that consists of the ability to identify thoughts and beliefs of others, and an affective side that allows an awareness of others' emotional states (Shamay-Tsoory et al., 2006). These proposals about what is included in the domain of ToM, more generally, can reflect how social and cognitive functioning can differ and concurrently affect each other. In the next section, I will describe how these two components are related, reporting neuroscientific evidences and results from researches on clinical populations.

1.3. Social Cognition and Executive Functions

Our abilities to reason, to plan, to remember, to shift our attention, represent a set of higher and lower demand skills that belong to the domain of cognition. Undoubtedly, these abilities serve as the basis of our functioning, playing an important part in maintaining and manipulating information, staying focused, initiating and planning actions, and problem solving. The cognitive literature also distinguishes a set of low order abilities, acknowledged as Executive functions (EFs). It is commonly agreed (Miyake et al., 2000; Diamond 2013) that the core of EFs is represented by: Inhibitory Control (IC), Working Memory (WM), and Cognitive Flexibility (CF). Previous literature debated about what it is included under the WM definition (Baddeley, 1986; Salthouse, 1991; Oberauer, et al., 2000). In this discussion WM refers to the ability to retain information and to perform mental operation. Inhibitory control involves the capacity to inhibit a response, to focus on a salient stimulus and suppressing other stimuli at the same time, and to control actions and thoughts. Cognitive Flexibility allows to shift from different tasks or mental set and promptly changing between one activities to another.

Given their nature, these cognitive abilities are important in our everyday life, such as remembering a phone number, stopping the car when the red lights appear. Nevertheless, these skills can be essential in social contexts. During a conversation we have to continuously update what the other person says, in order to respond in a correct manner. When we are watching the news and someone is talking in the same room, we are required to focus our attention on the person who is talking on the TV, and to inhibit the other voice in the room. However, if the news catches our interest, we have to switch our source of attention. With these simple everyday life examples, we can see how EFs are automatically used, allowing us to maintain a goal-direct behaviour and to “self-monitor” our own actions allowing us to interact in our environment in an appropriate way.

Wade et al., (2018) proposed three hypotheses on the relationship between ToM and EFs. The first one posits that EFs, in particular self-monitoring and inhibitory control, are essential for the development of ToM. The second hypothesis posits that ToM abilities are crucial for the development of EFs). In particular, children learn executive control through the mental representation of oneself and others, and consequently understand that the self and other are two separate identities (Lang & Perner, 2002; Perner, 1998). The third hypothesis, comes from clinical studies that have shown that impairments in ToM and EFs (e.g. in autism) share a common neural basis, the prefrontal cortex, and that damage to this area leads to deficits in both domains (Hughes & Ensor, 2007).

A long tradition of developmental research has demonstrated a robust relationship between the acquisition of EFs and improvements in ToM skills among young children, independent of age and IQ (e.g. Carlson, et al., 2004; Perner & Lang, 1999). However, the exact direction of this relationship remains under debate, with some researchers claiming that ToM is needed for EF (Carruthers & Smith, 1996; Perner, 1998; Perner & Lang, 1999, 2000), and others arguing that ToM requires EF (Russell, 1996, 1997; Pacherie, 1997). The specific EF skills that have been shown to be strongly correlated with ToM development are working memory (Keenan et al., 1998), inhibitory control (i.e. ignoring irrelevant information; Carlson et al., 2004), and cognitive flexibility (i.e. switching between different tasks; Hughes, 1998). These links make sense given the real-life examples above that show that successful social cognition requires one to hold in mind multiple perspectives (i.e. working memory), suppress irrelevant perspectives (i.e. inhibitory control), and switch between these two perspectives depending on context (i.e. cognitive flexibility).

Previous empirical work in adults has investigated whether the cognitive and social aspects correlate with each other, or if one moderates the other. Extensive literature has looked at perspective-taking (i.e. one of the components of ToM that indicates the ability to take the other persons' perspective), and EFs (Hartwright, et al., 2012; Lin, et al., 2010, Qureshi & Monk, 2018; Bradford, et al., 2015; Cane, et al., 2017), confirming that the ability to inhibit our point of view or shift from our perspective to another are involved in taking another person's perspective. Researchers have also examined the relationship between EFs and empathy. Empathy is defined as our ability to perceive others' emotional states and thoughts (i.e. cognitive empathy), as well as to share the other's emotional state (i.e. affective empathy). Gao et al., 2016, for example, reported that both aspects of empathy correlated with WM, however, in a recent metanalysis, Yan, et al., 2020, suggested that EFs are significantly related to cognitive empathy rather than with affective empathy.

Compelling evidence for a ToM-EF relationship comes from research in clinical populations (Uekermann, et al., 2008a; Zobel, et al., 2010). An impairment of one or more EFs can be detected in clinical syndromes that are characterised by deficits in social abilities, such as autism (Robinson, et al., 2009), conduct disorder (Happé & Frith, 1996), and ADHD (Willcutt, et al., 2005), as well as in neurodegenerative diseases like dementia (Poletti, et al., 2012), Parkinson (Verbaan, et al., 2007) and in neuropsychiatric disorders like schizophrenia (Braver, et al., 1999; Green, et al., 2015), obsessive-compulsive disorders (Watkins, 2005), depression (Uekermann, et al., 2008b; Zobel, et al., 2010). For example, in a study on patients with multiple sclerosis (MS), Ouellet et al., (2010) showed that patients with cognitive impairments experienced more difficulties in attributing mental states compared to other MS patients without or with moderate cognitive dysfunction. Interestingly, intervention studies have shown that ToM can be enhanced in autistic children by training ToM directly or indirectly via the underlying EFs (e.g. Fisher & Happé, 2005), suggesting that EFs are causally related to ToM, either as a prerequisite for ToM development or as crucial component of executive control of action. However, it remains under debate whether malfunctioning EFs cause difficulties in the social sphere, or vice versa, or whether there is an actual impairment in each of these two domains. Studies on clinical populations have tried to address this question, however, mixed results have been reported. Whereas some studies found a preserved cognitive functioning in conditions such as autism (Sucksmith, et al., 2013; Pellicano, 2007) and schizophrenia (Brüne, 2005a), others have provided evidence of intact social functioning in acquired brain injury (Aboulafia-Brakha, et al., 2011; Apperly, et al., 2009).

Further contributions to the debate on the relationship between social and cognitive processes come from brain imaging studies. Some studies report a common pattern of activation during tasks that involve EF and ToM over the frontal brain regions (see Figure 1. 1; e.g., Dohnel, et al., 2012; Hartwright, et al., 2012), including: the ventromedial Prefrontal Cortex (vmPFC - Hartwright et al., 2012; Shamay-Tsoory, et al., 2006), the inferior parietal lobule (IPL - Decety & Sommerville, 2003), the inferior frontal gyrus (IFG), and the tempo parietal junction (TPJ - has been identified crucial in disengaging and reorienting of attention; Apperly, et al., 2004; Corbetta, et al., 2008; Hartwright et al., 2012).

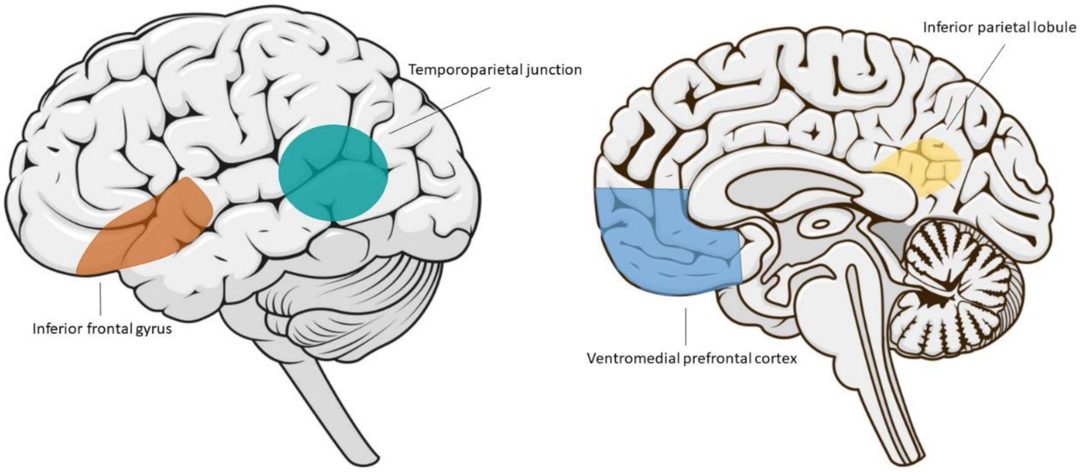


Figure 1.1: Common pattern of activation during EFs and ToM tasks.

Interestingly, brain stimulation studies with TMS over the dorsolateral PFC (dlPFC) have found improvement in ToM and impairment in EF (van den Heuvel, et al., 2013), suggesting an involvement of this brain area in both domains. A more recent study from Santiesteban et al., (2017) applied disruptive TMS during a perspective-taking task, and found that attentional control was reduced when participants had to judge the self or others' perspective.

Finally, a number of studies have examined how social cognition develops across the lifespan, which I will discuss later in this chapter, with a particular attention on whether social cognition abilities follow their own developmental trajectory or whether they are influenced by cognitive functioning. To provide appropriate context, I will first describe how cognitive abilities, and in particular EFs, progress across age.

1.4. Development of Executive Functions

Research has demonstrated a protracted period of EF development, which begins in early childhood and continues into young adulthood, with each sub-component of EF developing at its own rate (Diamond, 2002). Working memory and IC are the first cognitive components to appear in infancy. The ability to hold information in mind for a few seconds emerges around 6 months (Diamond, 1995; Reznick, et al., 2004), and at around 9 to 12 months infants can update the contents of their WM to track the position of an object from one place to another (Diamond, 1985). During this period, infants are also able to inhibit distractors in a similar way to adults (Fosco, et al., 2019). By the age of 6, working memory is quite well developed; children are able to solve complex tasks (Gathercole, et al., 2004), and are gradually able to hold more information in mind (Cowan, et al., 2010; Cowan, 2011). An improvement in the number of items recalled during a Visual Working Memory task seems to take place from 7 to13 years of age (Burnett Heyes, et al., 2012). Regarding IC, Carlson (2005) suggested that inhibition stabilizes around the early school years, however Johnstone et al., (2007) found that performance on a Stop-signal task (i.e. a task in which a prepotent answer has to be inhibit in response to a *stop-signal*) changed from ages 7 to 12 years old. Other researchers have supported this extended development of IC across childhood, reporting large improvements in a colour Stroop task from 10 to 15 years old (Prencipe, et al., 2011), in the Stop-signal task up to the age of 15 (Huizinga, et al., 2006), and on a Stroop task until the age of 21 (Huizinga, et al., 2006).

CF seems to appear later compared to WM and IC. For example, three years old children are able to categorize stimuli following a colour shape rule (Espy, 1997), but they commit errors in switching between these two categories. This difficulty could be explained as “attentional inertia”, a concept introduced by Kirkham, et al., 2003, children experience difficulty switching away from thinking about objects according to an initially relevant attribute. In their early childhood, children show prolonged response times (Kloo & Perner 2005; Chatham, et al., 2012) and longer eye fixations in tasks that involve switching abilities (Longman, et al., 2016). Research has shown improvements in task switching performance between 7 to 9 years of age (Davidson, et al., 2006), a reduction in error rates from 5 to 11 years old (Cohen, et al., 2001), and a decrease in RT between the ages of 7 and 15 (Huizinga et al., 2006).

Importantly, EFs continue to develop during adolescence (Diamond, 2002; Magar, et al., 2010). Neuroscience evidence supports a difference in the activation of brain areas across the lifespan. A common pattern of activation during cognitive functioning involves: frontal cortex areas including the dlPFC, and the anterior cingulate cortex (ACC), parietal areas, and subcortical structures, including the thalamus, caudate, putamen, the cerebellum (Niendam, et al., 2012). Children present enhanced activity of premotor regions and ventromedial areas (caudate nucleus and anterior insula; Scherf, et al., 2006), whereas in adolescence there is a greater activation of the frontal areas, and in particular, the dlPFC (Scherf, et al., 2006; Luna, et al., 2001), the inferior frontal, parietal, and anterior cingulate regions (Rubia, et al., 2006). These changes in activity coincide with key changes in EF abilities and are evidenced by MRI studies which have revealed a myelination of the prefrontal cortex and the connections between the prefrontal cortex and the subcortical areas, and also synaptic pruning loss of grey matter (O’Hare & Sowell 2008). These structural modifications in the prefrontal cortex, seem to reflect variations in cognitive functioning during adolescence (Blakemore, et al., 2006; Rubia, et al., 2000), and might lead to the mixed evidence about developmental trajectories. Overall, larger improvements in EFs occur during middle-childhood (5-11 years old; Romine & Reynolds, 2005), and become smaller between the ages of 11-14 years old (Hughes, 2011; Fosco, et al., 2019). CF and IC seem to reach a peak around the age of 12 (Bishop, et al., 2001; Van den Wildenberg & Van der Molen, 2004), whereas WM continues to advance until early adulthood (Gathercole, et al., 2004).

It is commonly agreed that EFs are subject to decline in healthy older age (e.g. Braver & West, 2008; Cepeda, et al., 2001; Lövdén, et al., 2010; Spieler, et al., 1996). Anatomically, older adults show a reduction in grey matter in fronto-temporal cortices (Ge, et al., 2002; Peters, 2006), and show more widespread brain activity compared to young people when managing tasks involving cognitive control; a consequence of reduced hemispheric lateralization (Cabeza, 2002). Moreover, older people show increased activity in prefrontal regions in tasks with low cognitive demands, but reduced activity in these same regions in tasks with high cognitive demands [“Compensation-Related Utilization of Neural Circuits

Hypothesis" (CRUNCH; Reuter-Lorenz & Cappell, 2008)]. It is generally agreed that working memory deteriorates in old age (Buckner & Louis, 2004; Klencklen, et al., 2017); older adults tend to experience more difficulty when under high WM load (Bennett, et al., 2013). Reed et al., (2014), proposed that age-related deficits in memory might be linked to attention and inhibitory control. Older adults experience difficulty suppressing irrelevant distractors (Gazzaley, et al., 2005; Zanto, et al., 2010) and employing selective attention (Engle & Kane, 2004), which leads to a failure to inhibit distractors and appropriately allocate attention in WM tasks (Davidson, et al., 2006; Solesio-Jofre, et al., 2012). Regarding cognitive flexibility, the literature has reported a general maintenance of this capacity (Wasylyshyn, et al., 2011; Brunsdon, et al., submitted), though some studies report difficulties keeping in mind the task goal, which may relate to WM difficulties (Cepeda, et al., 2001; Kray, et al., 2004). In sum, age-related changes in executive functions are thought to be relatively robust, but key differences in their magnitude and trajectory exist between different sub-components of EF, with some aspects of cognitive decline beginning from 20-30 years old, and decreasing at a faster rate with increasing age (Salthouse, 2009; Singh-Manoux et al., 2012).

1.5. Development of social cognition

An important concept when considering the development of social cognition is the "Social brain". Studies performed on monkeys by Brother in 1990 first proposed the idea of a "social brain" network, which includes all the brain areas that are involved in social cognition, including the amygdala, the orbital frontal cortex and the temporal cortex. Amodio and Frith (2006) later identified two more brain areas, the medial prefrontal cortex and the paracingulate cortex, that are activated in many in mental state tasks. Finally, Frith (2008; Frith & Frith, 2007) concluded the involvement of four specific brain areas in the "social brain" (see Figure 1. 2): the posterior superior temporal sulcus (pSTS) and the adjacent TPJ, the amygdala, the temporal poles, the medial prefrontal cortex (MPFC) and adjacent ACC, and the inferior frontal gyrus (IFG).

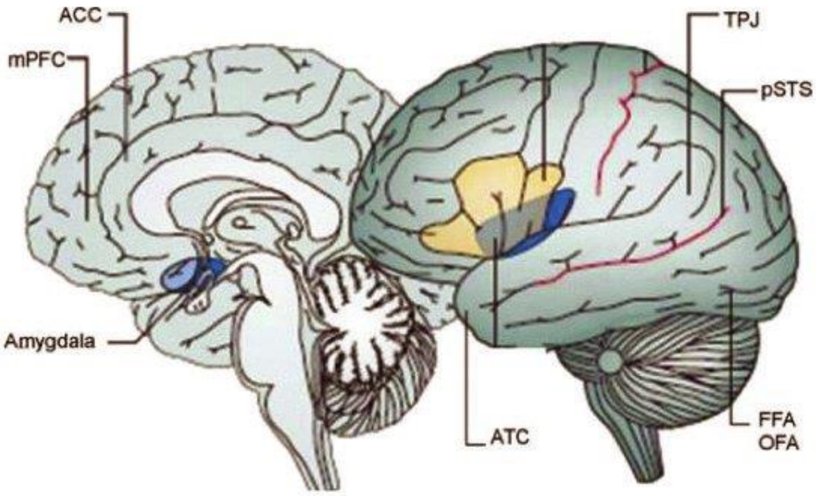


Figure 1.2: Social brain network (reproduced from Burnett et al., 2018).

In section 1.4. I outlined how cognition develops across the lifespan, and discussed how these variations are sustained by changes in the activity and structure of certain brain areas. This neurocognitive development corresponds with children's ability to pass increasingly complex tests of ToM. Around 18 months old, children are thought to have some implicit awareness of others' perspectives, but this is not consistently deployed in social situations (e.g. Buttelmann, et al., 2009; Kovács, et al., 2010; Onishi & Baillargeon, 2005; Rubio-Fernandez & Geurts, 2013; Senju, et al., 2011). Research conducted in childhood has focused on how children learn that others have their own minds, in other words the concept of ToM (Flavell, 2004). A largely adopted paradigm to investigate this understanding is the "false belief task" developed by Wimmer and Perner in 1983. In this paradigm, children are required to predict a character's behaviour by inferring their beliefs. For example, an object is moved from one location to another in the absence of the character, meaning that children have to inhibit their own knowledge about where the object really is and respond according to the character's more outdated knowledge. Typically, children younger than 3 years old fail this task (but see Rubio-Fernandez & Geurts, 2013), either because they cannot understand belief and others' minds (Wellman, et al., 2001) or they do not possess the higher cognitive skills that are necessary to resolve a conflict between reality and beliefs (Leslie, 2005). Overall, researchers have identified a key transitional phase in the development of ToM abilities around 4 years old (Banerjee, et al., 2011; Devine & Hughes, 2013), whereas later in childhood (around 7 or 8 years old) there seems to be an improvement in the comprehension of more complex beliefs (i.e. second order false beliefs; Perner & Wimmer, 1985). This development seems to be facilitated by the parallel progress of cognitive abilities (e.g. Carlson & Moses, 2001; Bock, et al., 2015).

Moreover, understanding intentions, thoughts, emotional states, recruits diverse components such as perspective-taking, emotion recognition, all abilities that develop across the childhood. Social cognitive abilities continue to develop throughout adolescence and well into our twenties (e.g. Blakemore, 2008; Burnett, et al., 2009; Dumontheil, et al., 2010; Symeonidou, et al., 2020; Vetter, et al., 2013), with adolescents appearing more egocentric (i.e. biased towards their own perspective) compared to young adults. In a longitudinal study Taylor et al. (2015), showed that social cognition abilities follow divergent trajectories between 17 and 19 years old (i.e. some functions improve, whereas others stabilize or decline), reflecting the different processes that are involved in the maturation of the brain (i.e. pruning, myelinization). Several fMRI studies have shown differences in the activation of the social brain network between adolescence and adulthood during social cognitive tasks (Blakemore, 2008, 2010). Specifically, during adolescence brain activity during social interaction tasks is higher in more posterior regions, such as the pSTS/TPJ (Blakemore, et al., 2007) and anterior temporal cortex (Burnett, et al., 2009). Further, the superior temporal regions, fundamental in ToM, mature later compared to those that sustain cognitive functioning (Apperly, et al., 2004). Decrements in grey matter volume across the frontal, temporal and parietal cortices during this stage of life has been linked to impulsivity, risk-taking behaviours and mood fluctuations (Casey, 2015; Anderson, et al., 2002; Steinberg, 2008). Indeed, children who reported higher control capacities aged between 3 to 11 years, including attention and persistence were more successful at school, less influenced to use drugs or smoke, and showed better physical and mental health (Moffitt, et al., 2011).

Early investigations of ToM in adulthood adopted many of the same measures developed for children (i.e. simple mind-reading tasks), leading to ceiling performance and high levels of accuracy. Researchers therefore developed new tasks to provide more sensitive indicators of ability and mechanisms in adults, by measuring accuracy and reaction times of responses (e.g. Apperly, et al., 2008; German & Hehman, 2006), or combining these methods with eye-tracking, ERPs or neuroimaging methods (e.g. Ferguson & Breheny, 2011, 2012; Ferguson, et al., 2015; Mahy, et al., 2014). Generally, these studies have shown that even in healthy adults, inferring other peoples’ mental states is more cognitively effortful than other inferences, and is subject to interference from our own point of view. For example, when required to take the perspective of another person or avatar, adults represent their egocentric perspective (Bradford, et al., 2019; Surtees & Apperly, 2012). Epley et al. (2004a) suggest that adults make judgements about others' perspective by first anchoring to their own point of view and consequentially adjusting it to fit with the other person; this adjustment requires cognitive effort. Qureshi (2009) pointed out that egocentric errors are a consequence of a failure to use information about the speaker's perspective, and they seem to be correlated with tasks in which participants have to hold in mind and select different responses (e.g. both cognitive capacities). Another study from German and Hehman (2006) revealed a correlation between the speed of inferring others' false beliefs and WM, IC and processing speed. Interestingly, different factors can influence adults' performance in ToM tasks, such as social relationships (Savitsky, et al., 2011), cognitive load (Cane, et al., 2017; German & Hehman, 2006), motivation and time pressure (Epley et al., 2004a; Cane, et al., 2017), as well as individual differences in EF (Brown-Schmidt, 2009b; Li, et al., 2010).

It is now generally acknowledged that healthy ageing is associated with increased difficulties in ToM (e.g. Henry, et al., 2014), however debate continues regarding the age that these declines first appear. Pardini and Nichelli (2009) suggest around 55 years of age, whereas Duval et al. (2011) suggest around 65 years of age. The existing literature has provided evidence for a decline in cognitive ToM that mirrors a general cognitive decline in old age (Bernstein, et al., 2011; Phillips, et al., 2011), however controversial results have been found on a potential connection between affective ToM and aging (Bottiroli, et al., 2016; Castelli, et al., 2010; Henry, et al., 2013; Mahy, et al., 2014b; Pardini & Nichelli, 2009). A very recent study from Yıldırım et al. (2020) found that older adults (51-80 years old) were impaired in short-term memory, processing speed, inhibition, and working memory, compared to their younger counterparts (18-28 years old), however, affective ToM abilities seemed to be spared. Importantly, other authors have highlighted other factors that can impair older peoples' performance in ToM tasks. For example, some studies have shown that middle-aged and older adults are *more* sensitive than younger adults to cues that facilitate social inferences (Hess, et al., 2005). Moreover, Duval et al. (2011) showed that age effects were not present when ToM was assessed subjectively, but clear age-related declines were visible when objective tasks were used to measure affective and cognitive components of ToM, and these were greatest for complex *versus* basic social inferences. This discrepancy between older adults' subjective experience of their mentalizing abilities compared to their objective performance on tasks that measure it highlights an impairment in metacognition.

To account for age effects on social cognition, a growing body of research has sought to examine whether the relationship between EFs and ToM changes across the lifespan (e.g. Leslie, 2005; Wade, et al., 2018; Sandoz, et al., 2014; Long, et al., 2018; Klindt, et al., 2017; Bailey & Henry, 2010; Charlton, et al., 2009; Phillips, et al., 2011). For example, a cross-sectional study by Klindt et al. (2017) examined mentalizing and cognitive abilities in participants ranging from 10 to 85 years old. The results of their "massive web poll" showed that low demand ToM skills decline at a slower pace than EFs, whose decline is slower than high demand ToM abilities; specifically, EFs contribute to high-level ToM after its complete maturation, suggesting that the cognitive architecture that sustains more advanced levels of mentalizing abilities changes across the lifespan. The authors also confirmed the interconnection between ToM and EFs, demonstrating that individual differences in EFs can predict ToM abilities. Overall an inverted U shape was found for both social and cognitive abilities, with a peak around twenty years and a slope starting around forty years. In another study, Duval et al. (2011) tested the relationship between affective and cognitive ToM and EFs among participants in three age groups: 21–34 years (young adults), 45–59 years (middle-aged), and 61–83 years (older group). The authors found that older adults performed more poorly when inferring cognitive and affective mental states compared to young and middle-aged adults. In particular, cognitive ToM scores were correlated with performance on EFs tasks. Interestingly, older adults perceived themselves as good mind readers despite the quantitative evidence to the contrary, which suggests that they experience a misperception about their objective decline. However, age differences in social inferences have been found to not be influenced by cognitive skills (e.g., Maylor, et al., 2002; Sullivan & Ruffman, 2004a).

To conclude, what clearly emerges in this section, is that literature has provided mixed findings on how social and cognitive abilities moderate each other. This might be due to different factors, such as targeting narrow age ranges or specific populations, testing small sample sizes, investigating isolated components of ToM or EF, or adopting mainly behavioural/response-based measures. These limitations leave open questions that we tried to address in the current work.

1.6. Training Social Cognition

Given the clear social consequences of delays or impairments in understanding others' minds, a number of ToM researchers have begun to explore whether young children can benefit from explicit training in ToM. Early work in this area has focused on training ToM skills directly (e.g. Appleton & Reddy, 1996; Slaughter & Gopnik, 1996), and has largely reported improvements in the trained aspect of ToM that also generalizes to other measures of ToM. In addition, a number of training studies have attempted to teach ToM skills to individuals with clinical disorders that are characterised by impaired social understanding (e.g. autism and schizophrenia), with some success.

Recently, a training program that involves engaging participants in conversation about mental states has been developed to enhance ToM (Lecce, et al., 2014). The authors found enhanced ToM abilities in children aged 9-10 years old in the Strange Stories task, and importantly, these improvements persisted two months after training. The same training protocol has been applied to healthy older adults aged 58-85 years old (Lecce, et al., 2015). The authors compared this 'social' training intervention based on describing mental states with a 'physical' training intervention (i.e. conversation focused on their physical sensations) and a social-contact group (i.e. participants had general conversations about aging). Results showed that ToM skills, measured with the Strange Stories task and a metarepresentational verbal task (i.e. a task used to investigate comprehension of mental states verbs), improved significantly more in the group who were trained using conversations about mental states. In a similar study, Cavallini et al. (2015) demonstrated that the benefits of social conversation training (*versus* physical conversation) in older adults held for both practised and transfer tasks. Together, these studies suggest that age-related decline in older adulthood might be mediated by training ToM directly, and that these abilities might generalise across tasks. However, it is not clear from these studies what mechanisms underlie these improvements, or specifically what aspects of the 'social' conversation produced improvements.

Other popular forms of training for ToM include those that use mindfulness training, or training protocols based on lectures and role play (e.g. Klimecki & Singer, 2012). Much of this work has targeted empathy. For example, a meta-analysis of interventions to reduce prejudice or promote positive intergroup attitudes in adolescents revealed moderate improvements in empathy after training (*d* = .3; Beelmann & Heinemann, 2014), with the strongest effects coming from interventions that involved direct interaction with another person, or social-cognitive training. Another meta-analysis highlighted that the efficacy of empathy training differs in different groups of individuals (e.g. health professionals and university students; Teding van Berkhout, & Malouff, 2016). Training interventions involving acting or role play have provided mixed evidence regarding benefits for empathy (Nettle, 2006; Goldstein, et al., 2009; Goldstein & Winner, 2010), but more consistent improvements in ToM (Marangoni, et al., 1995; Goldstein & Winner, 2010).

While some of these studies have demonstrated ToM improvements as a result of training (e.g. Bechi, et al., 2012; Wellman, et al., 2002), more recent work has highlighted significant problems with the generalizability of these effects to other ToM tasks and everyday mindreading (e.g. Begeer, et al., 2011), as well as poor maintenance of these trained skills over time (e.g. Williams et al., 2012). In this thesis, I will test the effectiveness of training ToM indirectly by training the relevant EFs, based on the developmental association between EF skills and ToM.

1.7. Structure of this thesis

In this thesis I will present three empirical chapters that investigate social cognitive abilities using large samples and a large battery of tasks, including behavioural measures, as well as mobile eye-tracking and EEG. The broad aim of this thesis is to investigate whether and how social cognitive abilities change across the lifespan, and whether we can enhance them through cognitive training.

In Chapter 2, I test a range of social cognitive abilities in a large sample of participants (N = 293) across three age groups: adolescents (aged 10-19 years old), young adults (aged 20 to 40 years old), and older adults (aged 60 to 80 years old). I aim to test some well-established effects of ToM, and test whether age moderates these abilities. This chapter tests the general hypothesis that social cognition abilities are enhanced among young adults compared to both adolescents and older adults. The methods and analyses were pre-registered and assessed three core aspects of social cognition (i.e. perspective-taking, empathy, and social attention) in six experiments, employing both lab-based and real-world measures of social cognition.

In Chapter 3, I examine the degree to which EF abilities can be enhanced through training in young adults (N = 160). I will present a pre-registered experiment that used a 21-day adaptive training procedure that directly compares the efficacy and generalisability across sub-components of EF using training programs that target WM, IC or CF *vs.* an active control group (i.e. was comparatively engaging and challenging, but did not train a specific EF). EFs will be assessed using a battery of tasks in a pre-post design. In particular, I will examine whether training transfers will include only the trained ability (e.g. transfer on WM tasks after WM training) or also to another EF ability (e.g. transfer on WM tasks after IC training).

In Chapter 4, I apply the training protocol developed in Chapter 3 to test whether social cognitive abilities can be enhanced indirectly through cognitive training, and whether these training effects differ with age. Social cognitive abilities will be assessed using the same six tasks from Chapter 2 and across the same three age groups: adolescents (aged 10-19 years old), young adults (aged 20 to 40 years old) and older adults (aged 60 to 80 years old).

CHAPTER 2: SOCIAL COGNITION IN ADOLESCENCE, YOUNG ADULTHOOD AND OLDER ADULTHOOD

Parts of this chapter have been submitted for publication:

De Lillo, M., Foley, R., Fysh, M., Stimson, A., Bradford, E., Woodrow-Hill, C., & Ferguson, H. J. (under review). Tracking developmental differences in real-world social attention across adolescence, young adulthood and older adulthood. *Nature Human Behaviour*.

De Lillo, M., Woodrow-Hill, C., Foley, R., Bradford, E., & Ferguson, H. J. (in prep). Empathy for others' physical and social pain in adolescence, young adulthood and older adulthood.

2.1. Introduction

As described in Chapter 1, some aspects of social cognition change with age (Henry, et al., 2013; Taylor, et al., 2013). In this chapter we sought to test a range of social cognitive abilities across three age groups: adolescents (aged 10-19 years old), young adults (aged 20 to 40 years old) and older adults (aged 60 to 80 years old). Our approach allowed us to address some of the limitations that have characterized earlier work. Firstly, most previous research on social cognition has focused on a very narrow age-range of 2-7 years old, when these skills are known to develop in typically developing children, or on clinical disorders that affect these skills (e.g. autism, schizophrenia). A growing body of research has emerged over the last few years, demonstrating that in fact social development continues through adolescence and well into our 20s (e.g. Blakemore, 2008), therefore it is possible that previous studies have overlooked key stages of ToM/EF development. In addition, the majority of studies that have assessed ToM in older age compare only two groups ('young' *vs*. 'old'), which does not allow comparison of social skills during development and decline. Secondly, most previous studies have used a single task to measure just one aspect of ToM, and typically record participants' explicit responses to questions that require an inference about a character's mental state (e.g. "where will X look for the Y?"). Such response-based methods offer limited insights in adults since they do not provide information on the processing steps that lead to a particular response; adults typically pass these explicit tasks but show some variation in performance when more sensitive measures are used. Here we use a broad range of state-of-the-art implicit methods (including eye-tracking and EEG) that measure success on several different aspects of social cognition, thus allowing us to capture distinct changes in each domain. In addition, many previous investigations of age-related changes in social cognition are limited by small sample sizes.

This chapter tests the general hypothesis that social cognition abilities are enhanced among young adults compared to both adolescents and older adults. Importantly, this research aims to address many of the limitations detailed above by including a large sample of participants (N = 293), pre-registering the methods and analyses, assessing three core aspects of social cognition (i.e. perspective-taking, empathy, and social attention) in six experiments, and employing both lab-based and real-world measures of social cognition. In the next sections, I will outline each of the three components of social cognition that are assessed here, and summarise what is currently known about age-related changes in these domains.

2.1.1. Perspective-taking

In social contexts we are continuously required to update our point of view in relation to others and to the surrounding environment. These processes are essential to communication and require perspective-taking to represent ourselves in space and to put ourselves in others' shoes (Keysar, et al., 2003). Representing ourselves is rather automatic and effortless. Through this egocentric point of view, we categorise everything around us relative to ourselves (e.g. front/behind, left/right). However, when communicating with others we must also consider *their* perspective and orientation in space to fully appreciate their mental state. Adopting the allocentric point of view is thought to be less automatic and more effortful. In this chapter we adopt three measures of perspective-taking to assess how these abilities change with age: mental state inferences, visual/spatial perspective-taking, and reference assignment (in the Director task).

Research suggests that the ability to make mental state inferences first emerges with intentionality (6-18 months), followed by desire (2nd year), then belief (4th year), and personality (6th–7th year) (Kalish & Shiverick, 2004; Wellman, et al., 2001; Wellman & Woolley, 1990; Woodward, 1998). Compared to young adults, older adults show impairments in identifying mental states (Henry, et al., 2013), and fail to take account intentions when judging moral permissibility (Moran, et al., 2012). Using a series of videos depicting everyday life situations, Lecce et al. (2019) showed that older adults were impaired in identifying and discriminating scenarios that involve mental states from those that do not require them. Younger participants tended to attribute mental states even when not requested, whereas older adults often reported mental states when they were not present or failed to detect mental states when they were. Adopting the same paradigm, Fossati et al. (2018) showed that adolescents performed better than adults, though both age groups committed errors in over identifying mental states. We assess mental state inferences using the 'hierarchy of social inferences' task developed by Malle and Holbrook (2012), in which young adult participants watched short videos and inferred mental states for the characters. Results revealed that judging intentions and goals were easier and faster than beliefs, followed by personality. We examine whether this hierarchy of social inferences changes across development.

The presence of a human avatar has been shown to influence perspective-taking (Samson, et al., 2010; Baker, et al., 2016; Furlanetto, et al., 2016; Todd, et al., 2017), suggesting that another person's view is automatically processed at an implicit level (Epley et al., 2004b; Samson et al., 2010; Surtees & Apperly, 2012). In fact, research has revealed interference from one's own point of view when adopting the avatar's perspective (egocentric intrusion; Epley, et al., 2004b; Keysar, et al.,

2003) as well as interference from the avatar's point of view when judging our own perspective (altercentric intrusion; Qureshi, et al., 2010; Todd, et al., 2017). Perspective-taking can be considered on several dimensions. Visual perspective-taking describes *whether* and *how* another person can see an object, and spatial perspective taking describes *where* an object is located in relation to another person. In addition, judging *what* another person can see (Level 1 perspective-taking) is distinct from judging *how* a person perceives something (Level 2 perspective-taking). While an egocentric intrusion occurs when judging what and how a person can see something (both Level 1 and 2), an altercentric intrusion seems to occur in Level 1 but not level 2 (Samson, et al., 2010; Surtees, et al., 2016; but see also Mattan, et al., 2017). Developmental studies showed that knowledge about *what* another person can see appears in the early childhood (14 months; Flavell, et al., 1981; Masangkay, et al., 1974; Sodian, et al., 2007), with 2-year-old children successfully passing tasks that tap this understanding (Moll & Tomasello, 2004). In contrast, the ability to comprehend *how* a person perceives something appears around 4 years of age (Flavell, et al., 1981) and continues to improve through early childhood (Frick, et al., 2014; Surtees & Apperly, 2012). Few studies have investigated perspective-taking in older age, however those that have revealed impairments in switching from their own perspective to the on-screen avatar's (altercentric) visual perspective (Martin, et al., 2019), and a preference to adopt a self-relevant perspective (Mattan, et al., 2017). We adopt Surtees et al. (2013a) paradigm that crosses the content of judgement (visual vs. spatial) with the type of judgement (level-1 type vs. level-2 type) to examine whether developmental changes differentially impact different components of perspective-taking.

Finally, perspective can be used to guide reference when interpreting instructions. In the so-called ‘Director task’, participants follow the instructions of an on-screen avatar to move objects around a grid. Importantly, some of those objects are occluded from the speaker’s but not the participant’s view, leading to a discrepancy in the two communicators’ perspectives (shared vs. privileged) and thus ambiguity in reference assignment. Typically, participants make more errors on privileged trials as they fail to adopt the speaker’s perspective and interpret reference according to their own egocentric perspective (e.g. Cane, et al., 2017; Ferguson & Cane 2017; Dumontheil, et al., 2010; Keysar, et al., 2000; Meyer, et al., 2015; Mills, et al., 2015; Santiesteban, et al., 2015). Research that has examined the developmental trajectory of perspective-taking in this task has shown that children (4 to 12 years) and adolescents (Choudhury, et al., 2006; Dumontheil, et al., 2010; Symeonidou, et al., 2016) commit more egocentric errors compared to young adults (Epley et al., 2004b), however no age differences emerged in terms of response time (Dumontheil et al., 2010). Currently no published research has used this task to examine how referential perspective-taking might change in older age.

2.1.2. Empathy for others’ pain

One of the most investigated components of social cognition is empathy. Empathy is defined as our ability to perceive and share the emotional states and thoughts of others (Davis, 1980). Empathy is sustained by a specific neural system located in the premotor cortex, called “mirror neurons” that activate when we observe actions performed by others (Gallese, et al., 1996; Rizzolati, et al., 2002; Hobson & Bishop, 2017). On this basis, Preston and de Waal (2002) proposed a perception-action model (PAM): seeing or imagining another person’s emotional state automatically activates a representation of this state in ourselves, generating a shared physiological response. Previous literature has revealed that observing others in pain activates the same brain areas as experiencing pain ourselves, in particular the AI, the ACC and the medial cingulate cortex (Craig & Craig, 2009; Craig, et al., 2003; Singer, et al., 2004; Hein, et al., 2010; Jackson, et al., 2005; Lamm, et al., 2010, 2011).

In this chapter we assess empathy for others' pain using behavioural ratings and EEG measures. A specific neural oscillation has been detected over the sensorimotor cortex when performing, observing or imaging actions (Hobson & Bishop, 2017; Pineda, 2005): mu rhythm, which includes alpha (8-13 Hz) and beta oscillations (13-35 Hz). Observing or imaging an action, or initiating a movement, generates a desynchronization of mu rhythm that leads to a decrease of its spectral power (Pineda, 2005). Previous studies on the experience of pain found that observing painful stimuli causes a suppression of the mu rhythm (Cheng, et al., 2008; Jackson, et al., 2005; Pineda 2005; Yang, et al., 2009), especially compared to non-painful stimuli (Perry, et al., 2010; Cheng, et al., 2008). Various factors seem to modulate the degree of pain, such as in-group vs out-group effects (Hein, et al., 2010), intensity of pain (Lamm, et al 2010), gender (Yang, et al., 2008), familiarity (Kross, et al., 2011), and type of stimulation (Cheng, et al., 2008).

Developmental studies on empathy have largely relied on behavioural measures and have revealed that the affective component of empathy first emerges in early childhood (Decety & Michalska, 2010; Cheng, et al., 2014), continues to develop throughout adolescence, then remains stable or increases though adulthood and older age (Sun, et al., 2018; Sze, et al., 2012; Ze, et al., 2014; Bailey, et al., 2020; Beadle & De La Vega, 2019). Older adults report higher state emotional empathy than younger adults in response to viewing empathy-eliciting film clips

(Richter & Kunzmann 2011), though younger participants rated needle-touches as more painful when directed to a younger versus older actor (Cao, et al., 2019). One study that used fMRI to compare empathy responses between adult age groups has shown that the neural response to others' pain reduces in older age (Chen, et al., 2014), while an EEG study found that young adults elicited increased mu activity for emotional compared to neutral stimuli, but older adults did not distinguish the two (Guay, et al., 2018).

In this chapter, we examine how empathic responses to others' pain changes from adolescence to young and older adulthood. We adopted the traditional paradigm whereby participants were presented with static images depicting hands and feet in painful (e.g. a needle piercing skin) and non-painful (e.g. a cotton bud pressing on skin) situations, and measured empathy via their ratings of imagined pain and mu suppression (alpha and low beta ranges). In addition, we designed stimuli to compare empathy for physical and social pain. Defined from Eisenberger (2012) as "An unpleasant experience that is associated with actual or potential damage to one's sense of social connection or social value (owing to social rejection, exclusion, negative social evaluation or loss)", very little research has examined empathy for social pain. One study employed the cyberball task, in which participants were either included or excluded in the game, and fMRI was recorded. Social exclusion elicited the same neural pattern as physical pain (i.e. dorsal anterior cingulate cortex (dACC) and anterior insula cortex (AIC); Krach, et al., 2011; Eisenberg & Lieberman, 2004; Eisenberger, 2011). This was further supported in an fMRI study that directly compared brain activity for vicarious physical (pictures of hands and feet in physically painful situations) and social (sketches that depict socially undesirable situations) pain (Krach, et al., 2015), which correlated with self-reports on the intensity of pain experienced. So far, only one study has used EEG to examine social pain in young children (aged 5-9 years old), and found that mu suppression was greater for videos depicting social pain compared to neutral events (Fraser, et al., 2020).

2.1.3. Social attention in the real-world

The majority of studies that have examined the development of social cognitive processing have been conducted in relatively tightly-controlled lab-based settings, in which individual participants merely observe other people in static images or dynamic videos; participants are not physically co-present in a social interaction (De Jaegher, et al., 2010; Schilbach, 2010). Although these lab-based designs have strengths in providing experimental control over stimuli, they are limited in ecological validity. Real-world social interaction and everyday use of social cognition is richer in detail and more nuanced than passively presented stimuli are able to convey (Foulsham & Kingstone, 2017; Risko, et al., 2012). Moreover, some studies have revealed inconsistencies in social behaviours when they are tested in a typical lab-setting versus an unconstrained real-world social interaction (e.g. Hayward, et al., 2017). It is therefore unknown whether the difficulties that adolescents and older adults show in lab-based social cognition tasks are magnified in real-world situations due to the complex and dynamic cues available, or whether the situational context scaffolds more successful use of social interaction processes.

To date, no research has examined the development of social attention while people actively participate in real-world interactive situations. Nevertheless, some insights can be gained from studies that have eye-tracked young and old adults while they watch videos of other people interacting. For example, Vicaria et al. (2015) and

Grainger et al. (2019) compared to older and younger adults' gaze as they watched videos of two people discussing a controversial topic, then judged the rapport between the two protagonists or their mental states. Both studies showed that older adults spent less time fixating on the protagonists' faces compared to young adults, and older adults were less able to correctly judge rapport or understand their mental states than young adults. Thus, in line with the impaired ToM seen in more tightly-controlled paradigms, older adults attended less to faces, and as a consequence experienced difficulty extracting social information about others. However, none of these studies have included an adolescent group to assess early development. Interestingly, adolescence is a period marked by particular sensitivity to the social environment (Peper & Dahl, 2013), most notably an attentional shift in social orientation from family members to their same-aged peers (Blakemore & Mills, 2014) and increased self-awareness (van den Bos, et al., 2011; Weil, et al., 2013). No studies have examined how social attention is allocated during adolescence, and whether this differs from adulthood.

In this chapter we explore whether social attention differs across development, testing whether social attention might be a mechanism to successful social interaction, and therefore a primary source of the impaired ToM seen in adolescents and older adults compared to young adults. We adapted two tasks that have examined social attention in real-world interactive situations, using mobile eye-tracking technology. Freeth et al., (2013; see also Vabalas & Freeth, 2016) monitored looking behaviours while young adult participants engaged in a semi-structured one-to-one conversation with an experimenter. Participants took turns with the experimenter to ask and respond to questions about general topics, such as plans for the weekend, and eye movements to the face, body and background were analysed while participants were speaking or listening. Results revealed a general preference to look at the experimenter's face, however social attention (i.e. fixations on the face) was higher when participants were listening compared to speaking, and attention to the non-social information (i.e. the background) was higher when speaking than listening. This pattern is interpreted as evidence that interlocutors found speaking more cognitively demanding than listening, and used gaze aversion as a means of reducing these processing costs to avoid distracting social information in the face (e.g. Barzy, et al., 2020; Doherty-Sneddon, et al., 2002; Doherty-Sneddon & Phelps 2005; Glenberg, et al., 1998). In another study, Foulsham, et al. (2011) recorded young adult participants' gaze while they walked through a busy university campus. Contrary to lab-based studies that have shown a human predisposition to preferentially attend to social information (Emery, 2000; Kingstone, 2009), the results from this naturalistic setting revealed that relatively few fixations were directed towards people (~22%) compared to other non-social information in the environment. This study therefore provides further evidence that social attention is reduced when participants are under additional cognitive demands (e.g. route-finding). In sum, the broad aim of this chapter is to investigate whether and how social cognition changes across the lifespan, assessing whether distinct patterns are seen across different sub-components of social cognition. Based on the literature discussed in Chapter 1 and above, we expected to find overall impaired social cognitive abilities in adolescents and older adults compared to young adults. Specific hypotheses for each task will be detailed in the Methods section.

2.2. Methods

All methodological procedures were pre-registered on the Open Science Framework (OSF) web pages (see osf.io/jgry4, osf.io/guf6k, osf.io/evb23, osf.io/qwz8m, osf.io/fnd8h, osf.io/za4nh).

2.2.1. Participants

A total of 293 participants, aged between 10-80 years old, were recruited for this study. Of this total sample, nine participants were excluded for having MoCA scores less than 26. This resulted in a final sample of 284 participants, divided into three age groups: 91 adolescents (aged 10-19 years), 104 younger adults (aged 20-40 years), and 89 older adults (aged 60-80 years). Participants were paid £50 for their time. All were native English speakers, had normal or corrected-to-normal vision, had no known neurological disorders, and no mental health or autism spectrum disorder diagnoses. Participants were recruited from a community sample in the local area of Kent, U.K., using a variety of recruitment strategies (e.g., newspaper adverts, local groups, word-of-mouth, Kent Child Development Unit). Sample size was pre-registered based on previous research, and constraints to complete the PhD. The Ethical Committee of the School of Psychology, University of Kent, U.K., approved the study.

Participant details, including mean age and gender balance for each of the three age groups, are presented in Table 2. 1. Socio-economic status (SES) was estimated by asking participants (if aged over 18) or parents of participants (if aged under 18) to report on their level of education, household income, and their occupation (job title and industry). To calculate an SES index, education level was coded on a scale of 1-6 (from No qualifications – Postgraduate Degree), and household income and occupational class were coded on a scale of 1-7. These three scores were summed to derive an SES index, with lower scores indicating lower SES. In addition, IQ was assessed using the Wechsler Abbreviated Scale of Intelligence – Second Edition (WASI; Wechsler, 1999), cognitive dysfunction was screened using the Montreal Cognitive Assessment (MoCA; Nasreddine, et al., 2005), and autistic traits were screened using the Autism Quotient-10 (AQ-10; Allison, et al., 2012).

Table 2.1: Participant characteristics by age group (mean values, with standard deviations in parenthesis).

	Adolescents	**Younger Adults**	**Older Adults**
N	91	104	89
Age (years)	14.6 (3.0)	27.0 (5.2)	67.9 (5.2)
Gender (F:M)	52:39	69:35	59:30
SES Index	10.1 (3.6)	10.6 (2.8)	11.2 (2.6)
Full Scale IQ	102.4 (10.2)	101.7 (13.3)	110.5 (11.1)
Verbal IQ	100.8 (9.0)	99.3 (10.8)	107.5 (12.4)
Perceptual Reasoning IQ	104.8 (11.8)	102.8 (12.1)	110.7 (13.3)
MoCA	27.9 (1.9)	27.7 (1.7)	27.3 (1.9)
AQ-10	2.7 (1.9)	3.2 (1.9)	2.6 (1.5)

2.2.2. Experimental tasks

2.2.2.1. Hierarchy of Social Inferences

We adopted the task developed by Malle and Holbrook (2012; Study 3) to examine the likelihood and speed with which people make mental state inferences about others, such their intentions, desires, beliefs and personality. Stimuli consisted of 42 videos that depicted people in everyday life situations and portrayed three classes of behaviours: goal-tailored (based on intentionality, e.g. a student riding his bicycle to university), trait-tailored (based on disposition, e.g. takes an orphan to the circus), and untailored (that elicited various inferences, e.g. a woman sweeping the floor). The videos ranged in length from 4 to 12s (M = 7s). Participants were instructed to make a judgement for each video about whether or not they detected a specific mental state in the main character's behaviour (note that some trials included a control or catch cue). A probe word, presented after each video, indicated the mental state to be inferred (see Table 2. 2).

Table 2.2: Inference probes and their meanings in the hierarchy of social inferences task

Probe word	***Question***
PERSONALITY	Did you detect a certain PERSONALITY characteristic the main actor has?
THEGOAL	Did you detect a certain GOAL the main actor has?
THINKING	Did you detect what the main actor was THINKING (was aware of, knew, saw, etc.) in this situation?
INTENTIONAL	Did you detect the actor INTENTIONALLY perform the behaviour?
ISMALE	Is the actor MALE?
DONOTRESPOND	When you see this cue, DO NOT answer

Each trial began with a central fixation cross for 1000ms, followed by the video, then a probe word appeared for 4000ms. Participants responded to the probe word using the keyboard. If they detected the probe social behaviour, they pressed the y key, which initiated a second screen asking them to explain their answer out loud (spoken responses were recorded using a microphone). If they did not detect the probe social behaviour, they pressed the n key, and the trial ended.

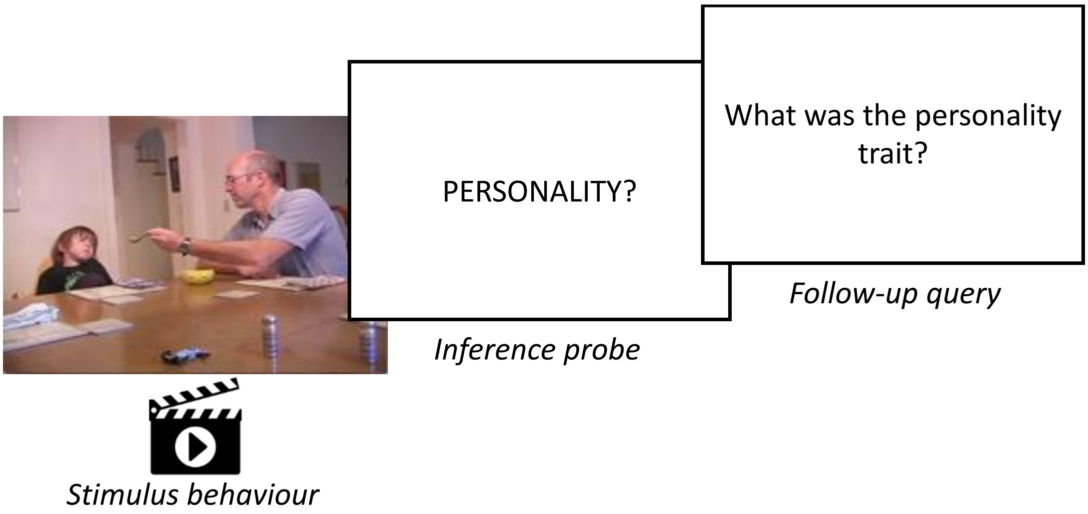


Figure 2.1: Event sequence for one trial in the hierarchy of social inferences task.

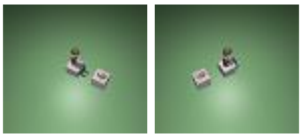
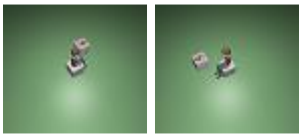
The experiment began with a familiarisation phase to ensure that participants understood the meaning of the inference probes. Participants were presented with each of the six probe words and had to say out loud the meaning, then were given the correct meaning again on screen. The main experiment consisted of a short practice block of eight trials, followed by 42 experimental trials divided into two blocks. Each of the four social behaviour probes appeared nine times during the experiment, while the catch and control probes appeared three times each. This task lasted 12 minutes on average.

Two dependent variables, likelihood of inferences and log reaction times, were analysed using a mixed ANOVA that crossed the between-subjects factor Age Group (adolescents *vs.* young adults *vs.* older adults) and the within-subjects factor Inference type (intention *vs.* desire *vs.* belief *vs.* personality). We predicted that overall participants would show higher likelihood of inferences and faster reaction times for intentions and desires (which should not differ), compared to beliefs, and finally personality. Regarding age effects, in our young adults' group we expected to replicate previous results from Malle and Holbrook in which participants showed a 'hierarchy of social inferences'. We predicted that participants would show higher likelihood of inferences and faster reaction times for intentions and desires (which should not differ), compared to beliefs, and finally personality. We also tested the hypothesis that this hierarchy of social inferences changes in adolescents and in old age, though make no directional predictions on these differences.

2.2.2.2. Visual/Spatial perspective-taking

Perspective-taking was assessed using the Visual/Spatial perspective-taking task developed by Surtees et al. (2013a). Stimuli showed a human avatar seated in a room with a cube that showed a number- 6 or 9. The cube could be positioned either in front, behind or to the side of the avatar, and the entire scene was presented to participants at different rotations (0°, 60°, 120°, 180°, 240°, 300°), as shown in Table 2. 3. Participants were instructed to take either the avatar’s visual or spatial perspective to judge WHAT (level 1) or HOW (level 2) they saw the cube.

Table 2.3: Difficulty levels used in the Visual Spatial perspective taking task.

	Level 1	**Level 2**
Visual	He CAN or CANNOT SEE the block 	He sees a 6 or 9 on the block 
Spatial	The block is IN FRONT or BEHIND him 	The block is to his LEFT or RIGHT 

In the level 1 visual perspective-taking condition, participants were asked to judge if the avatar can see or cannot see the cube. In the level 2 visual perspective-taking condition, participants had to indicate if the avatar can see the number 6 or 9. In the level 1 spatial perspective-taking condition participants responded if the cube was in front or behind the avatar, while in the level 2 spatial perspective-taking condition they had to judge if the cube was on the avatar’s left- or right-hand side. Thus, both level 1 conditions used the same set of images but gave different instructions to participants. In level 2 all the images contained instances in which the cube was visible to the avatar. For the two level 2 conditions and the level 2 visual perspective-taking condition the cube was either directly in front or directly behind the avatar, while in the level 2 spatial perspective-taking condition the cube was at a 45-degree angle from the avatar so a judgement could be made if the cube was on the left- or right-hand side of the avatar. Participants indicated their response to each trial using the up and down arrow keys on the keyboard.

Participants completed a total of four blocks (one for each condition), each consisting of 16 practice trials and 64 test trials. Angles 0° and 180° appeared 16 times each per block, and angles 60°, 120°, 240°, 300° appeared 8 times each; for analysis, the data from the 60° rotation was combined with data from the 300° rotation, and data from the 120° rotation was combined with data from the 240° rotation to include both clockwise and anticlockwise variations. The order of the four blocks, as well as the order of appearance of images within each block, was randomized. Each trial began with a central fixation cross for 600 ms, followed by 200ms blank screen and then stimulus presentation. Participants were given 5000ms to respond. If no response was recorded, the trial was coded as incorrect and next trial began automatically. Feedback on accuracy and response speed was provided during the practice blocks only. This task lasted 10 minutes on average.

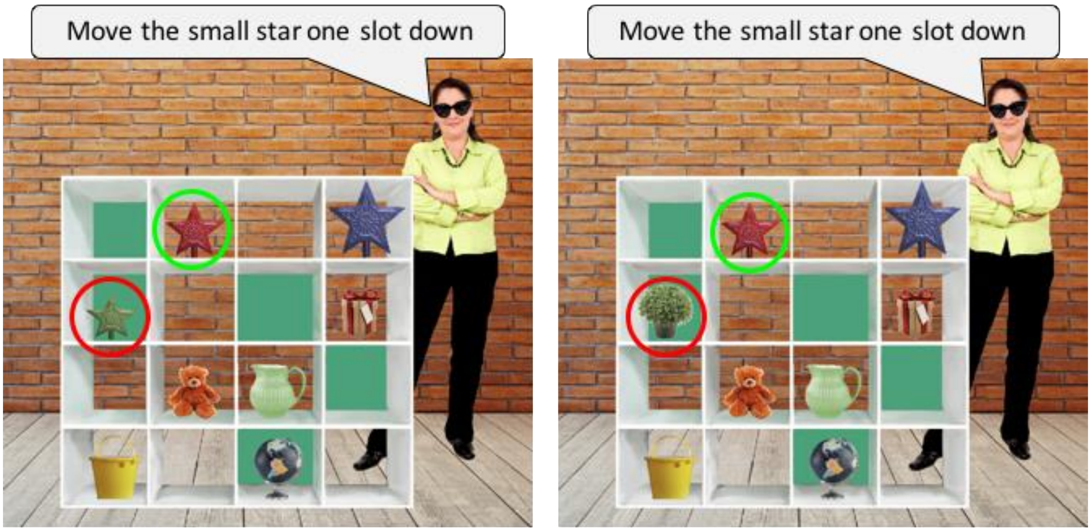
Two dependent variables, error rates and log reaction times, were analysed using a mixed ANOVA, crossing the between-subjects factor age Group (adolescents *vs.* young adults *vs.* older adults) with the within-subjects factors Angle (0° *vs.* 60° *vs.* 120° *vs.* 180°), Type (level-1 < level-2) and Content (visual < spatial). We expected to replicate previous studies (Surtees et al 2013a) in showing main effects for level (level 2 vs level 1), angle (increasing RTs with increasing angle), and content (visual vs spatial), as well as interactions for angle by level, and content by level. Regarding age differences, we predicted an improvement in response times through adolescence and a decline through old age in the ability to adopt another person's visual/spatial perspective. We predict that these impairments will be greatest for level 2 perspective-taking, and absent for level 1 perspective-taking.

2.2.2.3. Interactive reference assignment

To examine participants' ability to infer reference in conversation we used an avatar version of the referential communication 'Director' task (Keysar, et al., 2003). The task was delivered and controlled using SR Research Experiment Builder software (version 2.1.140). During the task, participants were presented with an image of a room containing a 4 x 4 gridded cupboard, creating 16 slots that could contain objects, and a female avatar (the 'director') standing to the rear right-hand side of the cupboard (see Figure 2. 2). Crucially, the backs of five slots (different for each trial) were covered with a green backing, so that only the participant could see the contents of these spaces, and the contents were occluded from the director's view. Eight objects were randomly placed within the grid slots, two of which were in occluded positions and six could be seen by both the director and the participant. Participants were asked to move objects around the grid following the avatar's verbal instructions.

In the Listener perspective condition, participants needed to take the director's perspective to select the mutually available object, and ignore a competitor object in an occluded slot that fitted the description (since it could not be seen by the director). Thus, participants were required to inhibit their own perspective to consider the correct object from the director's point of view. For instance, the participant could be asked to '*Move the small star one slot down*', where the grid contained three stars of different sizes, the smallest of which was occluded from the director. In this example, it would be correct for the participant to select the medium sized star, since this is the smallest star from the director's perspective. In the Shared perspective condition, the competitor object was replaced by a different (neutral) object that could not be mistaken for the object in the director's instruction. For instance, the participant could be asked to '*Move the small star one slot down*', where the grid contained only two mutually-available stars. Here, it would be correct to select the objectively smallest star, since this matched both the participant's and the director's perspective. Participants responded using the computer mouse to select an object, and drag it to the new location detailed in the verbal instruction. See Figure 2. 2 for a visual depiction of stimuli across Listener-Only and Shared-Perspective conditions.

The experiment included two practice trials and 24 experimental trials, of which 12 included a Listener perspective instruction and 12 included a Shared perspective instruction. Each trial included two instructions; one was a filler that referred to a specific item and did not involve perspective-taking (e.g. '*move the yellow bucket one slot up*). Filler instructions were not included in the analysis. The order of filler and critical instructions was counterbalanced across trials, and a new instruction was only given once participants had responded to the previous instruction. Audio instructions were presented through headphones, and participants were given 4000 ms before the first instruction to inspect the grid. This task lasted 10 minutes on average.



LISTENER PERSPECTIVE

SHARED PERSPECTIVE

Figure 2.2: Example stimuli used in the Interactive reference assignment task, showing Listener perspective and Shared perspective conditions.

Two dependent variables, error rates and log reaction times, were analysed using a mixed ANOVA that crossed the between-subjects factor Age Group (adolescents *vs.* young adults *vs.* older adults) with the within-subjects factor Condition (listener perspective *vs.* shared perspective). We predicted that participants overall will make more errors on listener only trials compared to shared trials. Moreover, we predicted that adolescents and older adults will make more egocentric errors on listener only trials compared to young adults.

Regarding age effect, we predict that participants will be faster to select the correct object when considering shared perspective versus listener only view. In addition, we expected reaction times to increase across the age groups (adolescents > young adults < older adults), and for the size of the listener-shared condition effect to be larger in the older adult group compared to the adolescent and young adult groups. To note that compared to previous literature, our study has adopted a shorter version of this task with a reduced number of target trials. This might affect the performance.

2.2.2.4. Empathy for physical and social pain

This task was based on Jackson et al. (2005), in which participants viewed images depicting others in painful or non-painful situations while brain activity was measured using EEG. Specifically, we compared the brain's response to images of hands and feet in physical and social pain, as an indicator of empathy for others. Stimuli were images and photographed events with real actors. Physical pain images depicted pain caused by pressure, thermal, sharp objects, etc, and social pain images depicted situations of embarrassment, grief, misery, etc (see Table 2. 4). An initial total of 162 images were tested and modified using a pre-test in which participants indicated the level of pain they thought the person in the picture was feeling, using a scale from 0 (no pain) to 100 (worst possible pain). Pain ratings were tested by a total of 89 students from the University of Kent using an online questionnaire platform (Qualtrics). Given the large number of images, the full set of 162 images was split into two sets of 81 images, and at least 41 participants rated each image. Images were presented one at a time, and participants rated the person's pain using a sliding scale. Using these ratings, we selected the final set of 40 images (see Appendix A for full set of images), which consisted of 10 physical pain images (M = 48.5, SD = 7.5), 10 physical no-pain images (M = 9.01, SD = 4.2), 10 social pain images (M = 51.2, SD = 11.7), and 10 social no-pain images (M = 11.1, SD = 5.1).

Table 2.4: Example stimuli used to depict physical and social pain and their corresponding no-pain images for the empathy for others' pain task.

		Content	
		Physical	**Social**
Type	**Pain**		
	No Pain		

The main experimental task consisted of 160 trials, 40 in each of the four conditions, and EEG activity was recorded throughout (for details see section 2.1.4). Each image measured 320x340 pixels, and was shown four times over experiment. Trials began with a central fixation cross for 500ms, followed by an image for 3000ms. On 25% of trials (i.e. once per image) a subsequent screen prompted participants to rate the level of pain that the person in the picture was feeling on a visual analogue scale from 0 (no pain) to 100 (worst possible pain); responses were made using the mouse. A blank screen was presented between trials using a variable inter-stimulus interval between 500 and 1500 ms to prevent expectancy effects on mu rhythm. Trials were presented in a randomized order, over four blocks. This task lasted 40 minutes on average, including EEG setup.

Three dependent variables were analysed. First, explicit pain ratings were analysed using a mixed ANOVA, crossing the between-subjects factor Age Group (adolescents *vs.* young adults *vs.* older adults) with the within-subjects factors Type (Painful *vs.* Non-painful) and Content (Physical *vs.* Social). Second, EEG data was analysed separately for mu/alpha (8-13Hz) and beta (13-35Hz) power, using a mixed ANOVA crossing the between-subjects factor Age Group (adolescents *vs.* young adults *vs.* older adults) with the within-subjects factors Type (Painful *vs.* Non-painful), Content (Physical *vs.* Social) and Electrode site (central *vs.* occipital).

We expected pain ratings to be higher for pictures that show painful situations compared to no-pain situations. We also predicted that this difference will be larger for physical pain compared to social pain. Regarding the EEG data, we expected to replicate previous studies in showing greater mu desynchronization for painful stimuli compared to no-pain stimuli (Perry, et al., 2010). Moreover, we expected that this pain-related desynchronization would be greater over the sensorimotor cortex than over occipital lobes, reflecting genuine mirror neutron activity rather than differences in arousal. We also predicted greater mu desynchronization for physical painful stimuli compared to social painful stimuli. Finally, in terms of age differences, we predicted that young adults would show greater mu desynchronization in response to painful stimuli compared to both adolescents and older adults.

2.2.2.5. Social attention in the real-world: face-to-face conversation

In this task, participants engaged in a semi-structured conversation with the experimenter while wearing the eye-tracking glasses, similar to Vabalas and Freeth (2016). The conversation tapped general topics, such as plans for the weekend or hobbies. In the first part, the experimenter asked four questions (e.g. “Tell me about some things that you did last weekend and some things that you plan to do next weekend” or “Describe a few things you consider to be typically English and a few things you consider to be typically American”) that were designed to prompt the participant to speak for approximately 30s (see Appendix B for the full set of questions); this was defined as the *Speaking* phase. In the second part, the participant and experimenter switched roles, and the participant now asked the same questions to the experimenter and listened to their answers; this was defined as the *Listening* phase. The Speaking phase was always followed by the Listening phase. The experimenter sat in a chair opposite the participant, approximately one meter away, and looked directly at the participant while speaking and listening. Verbal responses were recorded through a microphone integrated into the glasses. This task lasted ~10 minutes on average.

As an additional measure of social attention, we displayed three posters directly behind the experimenter (see Figure 2. 3). Two of these posters depicted social scenes- a group of young adults either with averted gaze (i.e. looking at each other) or shared gaze (i.e. looking out of the image towards the participant), and one poster depicted a non-social scene (i.e. a scene from the local area with no people). The position of the posters was randomized across the participants, and two sets of images were used for each condition (see Appendix C).

Figure 2.3: (left panel) A typical view seen by a participant during the face to face conversation task; (right panel) example definitions of the three AoIs used to analyse fixations in this study.

Two separate analyses were conducted on the proportion of fixations during this conversation task. The main analysis used a mixed ANOVA, crossing the between-subjects factor age Group (adolescents *vs.* young adults *vs.* older adults) with the within-subjects factors Condition (speaking *vs.* listening) and AoI (face *vs.* body *vs.* background). Eye movements towards the background posters were analysed using a mixed ANOVA, crossing the between-subjects factor age Group (adolescents *vs.* young adults *vs.* older adults) with the within-subjects factors Condition (speaking *vs.* listening) and AoI (shared gaze *vs.* averted gaze *vs.* neutral). We expected to replicate the basic effects from Freeth et al. (2013); overall participants would preferentially attend to social content in the environment, and this would be modulated by phase of the conversation (i.e. more looks to the experimenter’s face while listening, and to the background while speaking). We also predicted that adolescents and older adults would make fewer fixations towards their social partner compared to young adults. Finally, based on studies showing a bias to look at people (Birmingham, et al., 2009), we predicted that overall participants would make more fixations to the background poster depicting social scenes compared to the poster depicting a non-social scene.

2.2.2.6. Social attention in the real-world: Navigating an environment

This task adapted the real-world navigation task used in Foulsham et al. (2011). Participants were asked to complete a short independent task, to walk from the lab to College reception to collect a leaflet, and walk back. They were told they could take the route of their choice, and were given a map of the building to ensure that participants walked through similar environments (see Figure 2. 4). This task involved a walk of 5-10 mins inside a building environment that featured objects, signs, and other people. The experimenter followed participants from a distance.

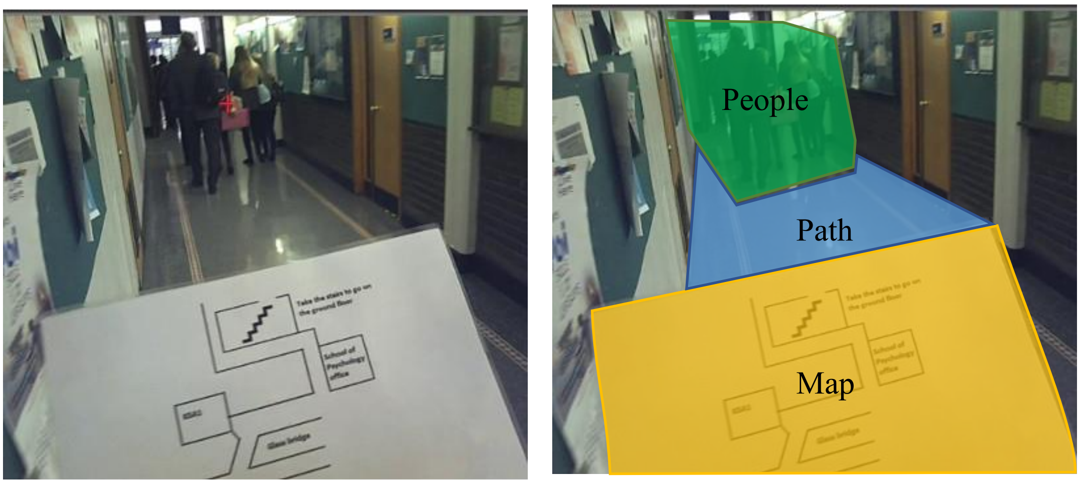


Figure 2.4: (left panel) An example view of the environment seen by a participant during navigation task; (right panel) example definitions of the AoIs used to analyse fixations in this study.

Analysis was conducted on the proportion of fixations during this navigation task, using a mixed ANOVA that crossed the between-subjects factor age Group (adolescents *vs.* young adults *vs.* older adults) with the within-subjects factor AoI (map *vs.* objects *vs.* path *vs.* people). We expected to replicate the basic effects from

Foulsham et al. (2011); overall participants would make more fixations on non-social content (i.e. objects, map, path) than social content (i.e. people) in the environment. We also predicted that older adults would make fewer fixations towards people compared to adolescents and young adults, and instead would look more at non-social content.

2.2.3. Procedure

Participants completed the six social cognition tasks in a single testing session in a quiet laboratory at the University of Kent. The order of tasks was counterbalanced across participants, and was run as part of a larger task battery (that included the cognitive assessment tasks described in Chapter 3). The entire testing session lasted approximately 3.5 hours, including breaks when needed.

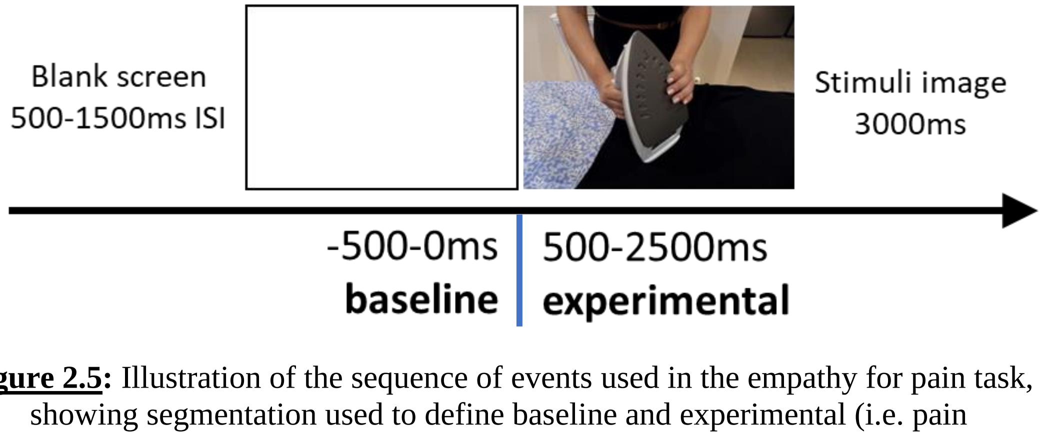
2.2.3.1. EEG recording and analysis

Electroencephalographic (EEG) activity was recorded during the empathy for physical and social pain task from 30 active electrodes using a Brain Vision Quickamp amplifier system with an ActiCap cap referenced to FCz. Vertical electro-oculogram (VEOG) activity was recorded from one extra electrode (below right eye), and horizontal electro-oculogram (HEOG) activity was recorded from one extra electrode (to the left of the left eye). EEG and EOG recordings were sampled at 1000 Hz, and electrode impedance was kept below 10kΩ.

Prior to segmentation, a vertical ocular calculation was applied (1*Fp2+(-1*VEOG)). All data were re-referenced to a common average reference. EEG and EOG activity was band-pass filtered (0.1-70 Hz, notch filter at 50Hz). Data were visually inspected for noisy sections or channels, and for other general artefacts.

EEG activity containing blinks was corrected using a semi-automatic ocular ICA correction approach (Brain Vision Analyzer 2.1). An average of 3.89 ICA components were removed per individual dataset.

EEG data was time-locked to the onset of each stimulus image, and data was segmented into a 500ms baseline period (-500-0ms from stimulus onset) and a 2s pain judgement period (500 - 2500ms from stimulus onset), as shown in Figure 2. 5. Semi-automatic artefact detection software (Brain Vision Analyzer 2.1) was run, to identify and discard segments with non-ocular artefacts (drifts, channel blockings, EEG activity exceeding ± 50µV). A fast-fourier transformation, with 10% Hanning window, was then applied to each segment, and the signal was averaged for each condition and electrode. In the baseline period, there was an overall data loss of 5% for the physical pain trials, 5% for the physical no pain trials, 5% for the social pain trials, and 5% for the social no pain trials, with an average of 37.98 (out of 40) baseline trial segments retained per participant. In the pain judgement period, there was an overall data loss of 8% for the physical pain trials, 9% for the physical no pain trials, 9% for the social pain trials, and 8% for the social no pain trials, with an average of 36.61 (out of 40) trial segments retained per participant.



<u>Figure 2.5</u>: Illustration of the sequence of events used in the empathy for pain task, showing segmentation used to define baseline and experimental (i.e. pain judgement) periods.

The average mu/alpha (8-13Hz) and beta (13-35Hz) power for each condition was calculated for the electrodes of interest over the central (C3, Cz, C4) and occipital electrodes (O1, Oz, O2), as shown in Figure 2. 6. This allowed us to test whether changes in mu and beta desynchronization over central sites were distinct from alpha and beta desynchronization over occipital sites. A measure of the percentage change in power was calculated for each experimental condition (physical pain, physical no pain, social pain trials, and social no pain) relative to the baseline period in that same condition for each electrode of interest in both alpha and beta bands, using the formula: log10(experimental/baseline). Data from electrodes C3, Cz and C4 was averaged for the central electrode site, and data from electrodes O1, Oz and O2 was averaged for the occipital electrode site. Negative values indicate mu/alpha and beta suppression.

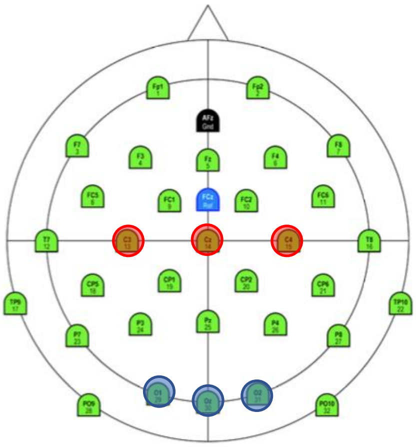


Figure 2.6: Central electrodes and occipital electrodes, are indicated with red and blue colours, respectively.

2.2.3.2. Eye-tracking recording and analysis

SMI mobile eye-tracking glasses were used to record real-life eye movements. A front-facing camera on the glasses recorded a video of the scene (field of view: 60° horizontal, 46° vertical; resolution: 1280 x 960pixels), as seen by participants, and binocular eye movements around this scene were recorded at a sample rate of 60Hz (with 0.5° accuracy). Corrective lenses of the appropriate prescription could be attached to the eye-tracking glasses if necessary. Before each real-world task, participants were fitted with the eye-tracking glasses, the experimenter ensured that they were comfortable, and participants completed a 3-point calibration and validation procedure.

SMI BeGaze analysis software (3.7.59) was used to prepare fixation data for analysis. For the face to face conversation task, fixations during the verbal responses were assigned to one of six areas of interest (AoIs): the experimenter's face, body, background, shared gaze poster, averted gaze poster, neutral poster. The background AoI was defined as any area in the scene except for the experimenter, and included fixations to the posters for the main analysis. For the navigation task, fixations were assigned to one of four areas of interest (AOIs): path, map, objects, and people.

2.3. Results

All analysis procedures were pre-registered, and the full datasets and analysis scripts are available on the Open Science Framework web pages (see osf.io/jgry4, osf.io/guf6k, osf.io/evb23, osf.io/qwz8m, osf.io/fnd8h, osf.io/za4nh). All statistical analyses were conducted in R version 3.6.1.

2.3.1. Hierarchy of Social Inferences

As in Malle and Holbrook (2012), likelihood was calculated as the percentage of trials on which participants responded "yes" that they detected a social behaviour, and reaction times were calculated for trials with a "yes" response. Outliers were excluded from the analysis of reaction times if they fell more than 2.5 standard deviations from the age group's mean response time and if they were lower than

400ms. The final sample in this task was 282 (90 adolescents, 103 young adults, 89 older adults). Two participants were excluded due to failure of the program. A 3 x 4 mixed design ANOVA was used to analyse the likelihood of inferences and log-transformed reaction time data, crossing the between-subjects variable age Group (adolescents *vs.* young adults *vs.* older adults) with the within-subjects variable Inference (Intention, Desire, Belief, Personality). Full statistical effects are reported for likelihood in Table 2.5, and for reaction times in Table 2.6.

Table 2.5: Statistical effects for Likelihood in the Hierarchy of social inferences task. Asterisks show significance of effects, where ** $p < .01$; *** $p < .001$.

	df	*F*	*p*	η_p^2
Age Group	(2, 279)	16.74	< .001***	.11
Inference	(3, 837)	141.86	< .001***	.34
Age Group x Inference	(6, 837)	2.72	.01**	.02

Analysis of likelihood revealed significant main effects for both variables. The main effect of Age Group showed that overall, adolescents (*M* = 71%) were less likely to make social inferences compared to young adults (*M* = 79%; *t*(191) = 3.43, *p* < .001), but older adults were more likely make social inferences (*M* = 84%) compared to young adults (*t*(190) = 2.77, *p* = .006). The main effect of Inference was further explored following our pre-registered contrasts, which showed that participants were more likely to more likely to infer a desire (*M* = 88%) compared to an intention (*M* = 85%; *t*(281) = 2.83, *p* = .005), more likely to infer an intention compared to a belief (*M* = 76%; *t*(281) = 7.34, *p* < .001), more likely to infer a desire compared to a belief (*t*(281) = 10.10, *p* < .001), and more likely to infer a belief compared to personality (*M* = 63%; *t*(281) = 8.77, *p* < .001).

The Group x Inference interaction was also significant. Follow-up analyses used 1-ANOVAs to test whether the likelihood of making each social inference differed between the three age groups. Results revealed that age Group modulated the likelihood of making inferences about the main character's intentions, *F* (2, 279) = 14.68, $p < .001$, $\eta_p^2 = .10$, desires, *F* (2, 279) = 10.4, $p < .001$, $\eta_p^2 = .07$, and beliefs, *F* (2, 279) = 16.91, $p < .001$, $\eta_p^2 = .11$, but not inferences about personality, *F* (2, 279) = 2.08, $p = .13$. Planned comparisons between age groups revealed that adolescents were less likely than young adults to infer social behaviour in the videos [intentions, *t*(191) = 3.33, $p = .001$; desires, *t*(191) = 3.67, $p < .001$; beliefs, *t* (191) = 2.84, $p = .005$]. In contrast, older adults were more likely than young adults to infer social behaviour in the videos [intentions, *t*(190) = 2.19, $p = .03$; beliefs, *t*(190) = 3.42, $p < .001$]. Contrasts between Inference types revealed the same significant patterns for all age groups (all contrasts, *p*s < .04, except when comparing intentions and desires in the old group, $p = .63$), which suggests that the basic hierarchy of social inferences does not change with age.

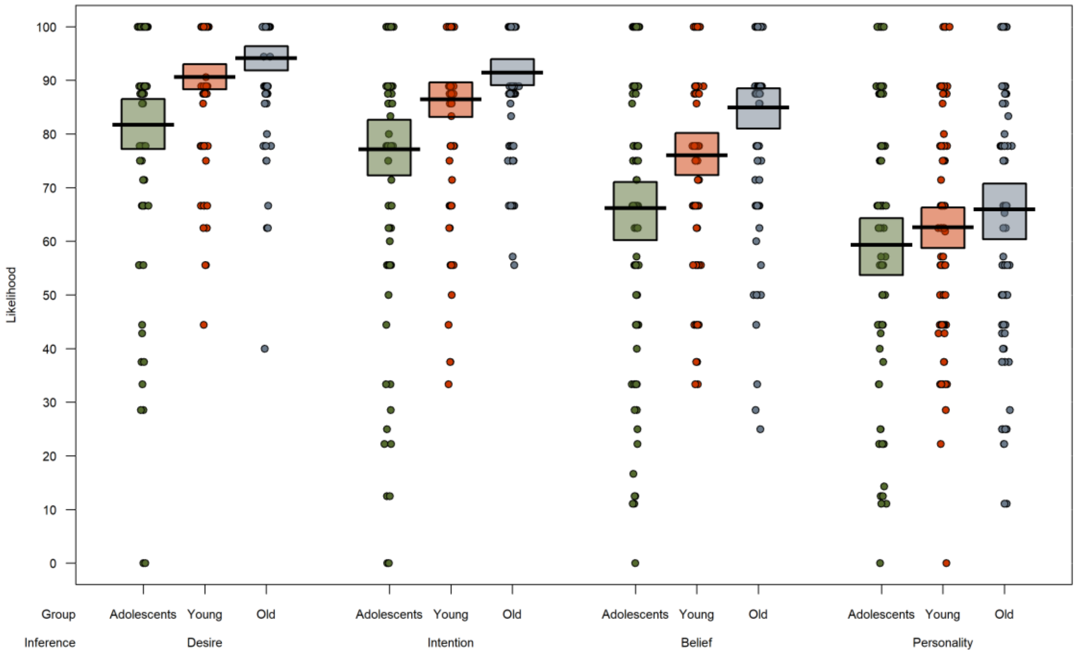


Figure 2.7: The likelihood of making social inferences in each age group in the Hierarchy of social inferences task. The plots show raw data points, a horizontal line reflecting the condition mean, and a rectangle representing the Bayesian highest density interval.

Table 2.6: Statistical effects for reaction times in the Hierarchy of social inferences task. Asterisks show significance of effects, where * $p < .05$; *** $p < .001$.

	df	*F*	*p*	η_p^2
Age Group	(2, 279)	29.21	< .001***	.17
Inference	(3, 837)	11.87	< .001***	.04
Age Group x Inference	(6, 837)	2.14	.05*	.02

Results on reaction times once again revealed significant main effects for both variables. The main effect of Age group showed that older adults were slower to make social inferences (*M* = 2116ms) compared to young adults (*M* = 1695ms; $t(190) = 8.00$, $p < .001$), but young adults and adolescents (*M* = 1773ms) did not differ, $t(191) = 1.93$, $p = .06$). The main effect of Inference showed that participants were equally fast at inferring a desire (*M* = 1771ms) and an intention (*M* = 1783ms; $t(281) = .53$, $p = .60$), faster to infer an intention compared to a belief (*M* = 1860ms; $t(281) = 3.64$, $p < .001$), faster to infer a desire compared to a belief ($t(281) = 3.20$, $p = .002$), and did not differ for inferring a belief and personality (*M* = 1899ms; $t(281) = 1.69$, $p = .09$).

The Age Group x Inference interaction was also significant. Follow-up analyses used 1-way ANOVAs to test whether the speed of responses of making each social inference differed between the three age groups. Results revealed that Age Group modulated the response times of making inferences about the main character's intentions, $F(2, 279) = 16.49$, $p < .001$, $\eta_p^2 = .11$, desires, $F(2, 279) = 34.73$, $p < .001$, $\eta_p^2 = .20$, beliefs, $F(2, 279) = 15.85$, $p < .001$, $\eta_p^2 = .10$, and personality, $F(2, 279) = 20.19$, $p < .001$, $\eta_p^2 = .13$. Planned comparisons between age groups revealed that adolescents were less likely than young adults to infer social behaviour in the videos [intentions, $t(191) = 2.14$, $p = .03$; desires, $t(191) = 2.01$, $p = .04$]. In contrast, older adults were more likely than young adults to infer social behaviour in the videos [intentions, $t(190) = 6.34$, $p < .001$; desires, $t(190) = 8.58$, $p < .001$; beliefs, $t(190) = 5.76$, $p < .001$; personality, $t(190) = 6.21$, $p < .001$]. The other contrasts did not reach significance. Contrasts between inference types were conducted separately per each age group. Results revealed that older participants responded significantly faster when judging intentions (M = 2023ms) compared to desires (M = 2139ms, $t(88) = 3.33$, $p = .001$) and beliefs (M = 2119ms, $t(88) = 2.67$, $p = .009$); younger, instead responded significantly slower when judging beliefs (M = 1769ms) compared to intentions (M = 1649ms, $t(102) = 2.88$, $p = .005$), and desires (M = 1642ms, $t(102) = 3.90$, $p < .001$); whereas adolescents responded significantly faster when judging desires (M = 1737ms) compared to beliefs (M = 1781ms, $t(89) = 2.13$, $p = .04$).

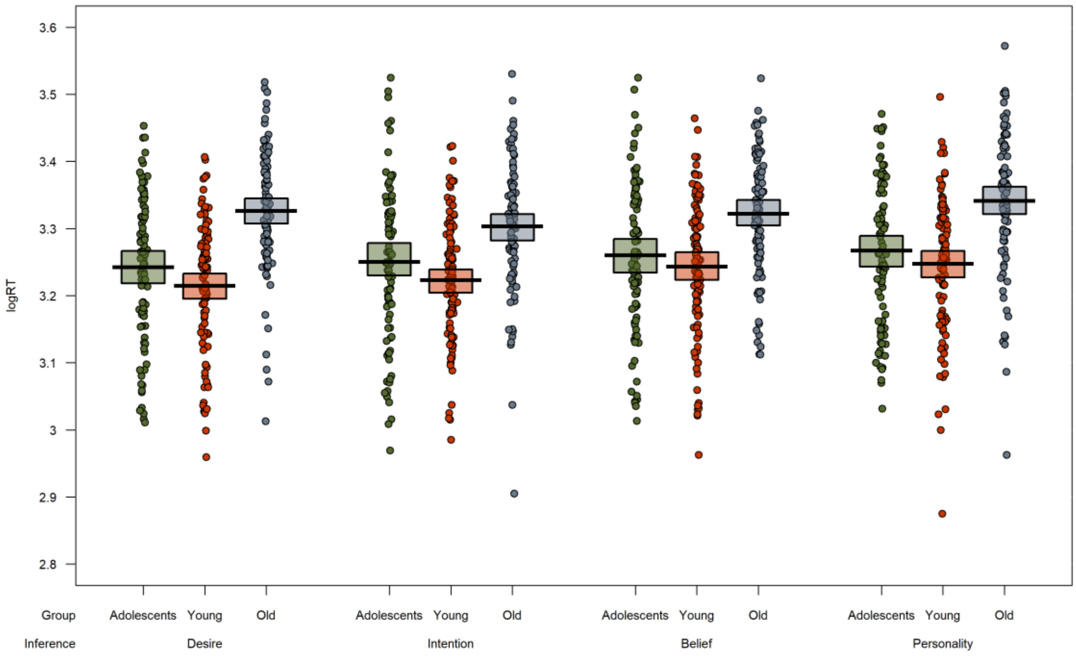


Figure 2.8: The log reaction time in each age group in the Hierarchy of social inferences task. The plots show raw data points, a horizontal line reflecting the condition mean, and a rectangle representing the Bayesian highest density interval.

2.3.2. Visual/Spatial perspective-taking

As in Surtees et al., (2013a), only matching trials were included in our analyses, and reaction times were calculated only on correct trials. Outliers were excluded from the analysis of reaction times if they fell more than 2.5 standard deviations from the age group's mean response time. The final sample in this task was 282 (90 adolescents, 104 young adults, 88 older adults); one participant from the adolescent group was excluded from the analysis due to accuracy lower than 50% and one due to computer failure. A 3 x 4 x 2 x 2 mixed design ANOVA was used for analysis of log-transformed reaction time data, crossing the between-subjects variable age Group (adolescents *vs.* young adults *vs.* older adults), with within-subjects variables Angle (0°, 60°, 120°, 180°), Level (1, 2) and Content (Visual, Spatial). Note that due to space constraints, we only report in the text follow-up analyses for interactions that were predicted/planned in the pre-registration, or involved age Group. Full statistical effects are reported in Table 2. 7. Data on accuracy are available in the Appendix D.

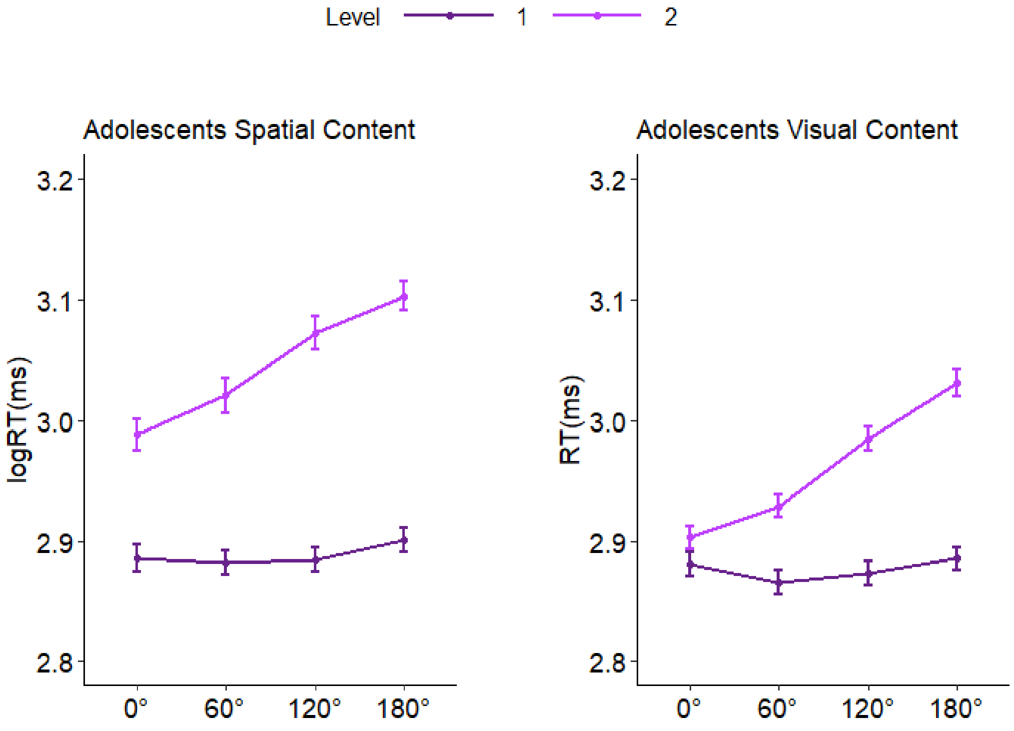
Table 2.7: Statistical effects for reaction times in the visuo-spatial perspective-taking task. Asterisks show significance of effects, where * $p < .05$; *** $p < .001$.

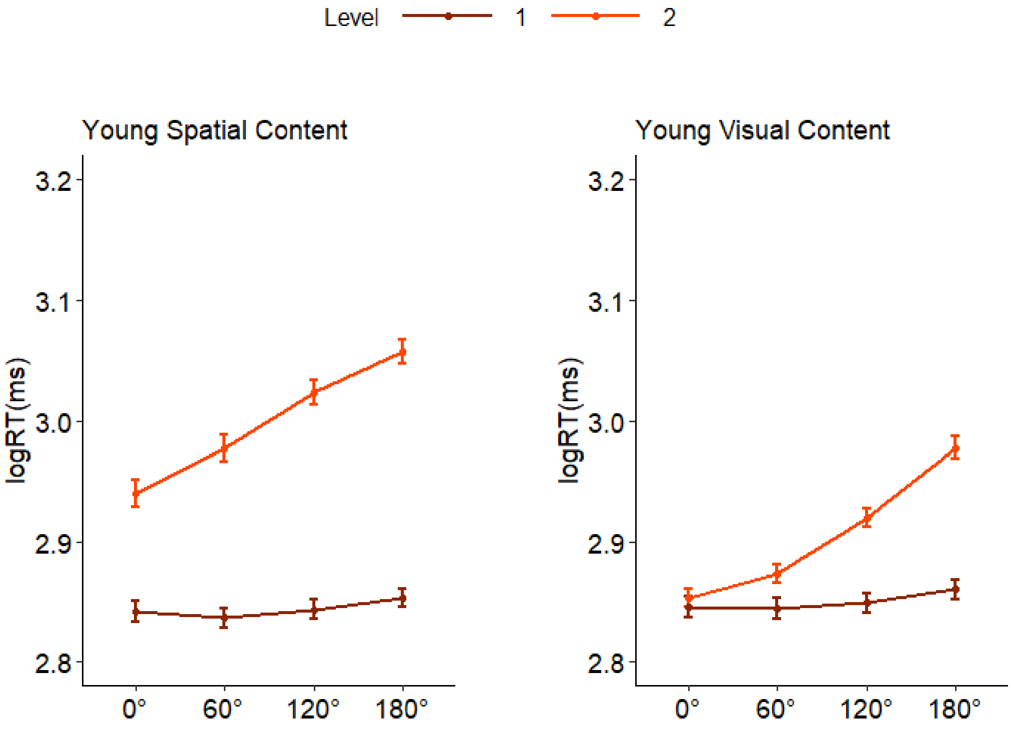
	df	*F*	*P*	η_p^2
Age Group	(2, 279)	69.40	< .001***	.33
Content	(1, 279)	204.80	< .001***	.42
Angle	(3, 837)	395.97	< .001***	.59
Level	(1, 279)	662.13	< .001***	.70
Age Group x Content	(2, 279)	0.88	.42	< .01
Age Group x Angle	(6, 837)	8.10	< .001***	.06
Age Group x Level	(2, 279)	4.80	.009	.03
Content x Angle	(3, 837)	7.06	< .001***	.03
Content x Level	(1, 279)	141.23	< .001***	.34
Angle x Level	(3, 837)	282.19	< .001***	.50
Age Group x Content x Angle	(6, 837)	1.77	.10	.01
Age Group x Content x Level	(2, 279)	2.23	.11	.02
Age Group x Angle x Level	(6, 837)	2.22	.04*	.02
Content x Angle x Level	(3, 837)	7.94	< .001***	.03
Age Group x Content x Angle x Level	(6, 837)	1.53	.17	.01

Results revealed significant main effects for all variables. Overall participants responded slower when making spatial compared to visual judgements (991ms *vs.* 870ms), and when adopting a level 2 perspective compared to a level 1 perspective (1049ms *vs.* 812ms). The main effect of Angle showed that response times increased as the angle of rotation increased ($M_{0°} = 875$; $M_{60°} = 892$; $M_{120°} = 943$; $M_{180°} = 1012$; all contrasts $p < .001$). The main effect of Age group showed that older adults spent longer judging the avatar's perspective (1012ms) compared to both adolescents (917ms) and young adults (821ms), and adolescents were slower compared to young adults (all contrasts $p < .001$).

As expected, the Content x Level interaction was significant. Follow-up analyses tested the effect of content separately for each level, and showed that response times were longer when judging the avatar's spatial *vs.* visual perspective for both level 1, $t(281) = 2.07$, $p = .04$, and level 2 judgements, $t(281) = 15.34$, $p < .001$. The Angle x Level interaction was also significant. Follow-up analyses tested the effect of angle separately for each level, and showed a significant effect of Angle on both level 1, $F(3, 843) = 18.92$, $p < .001$, $\eta_p^2 = .06$, and level 2 perspective-taking, $F(3, 843) = 545.50$, $p < .001$, $\eta_p^2 = .66$. For Level 1 judgements a significant difference only emerged on angle contrasts 0° *vs.* 60°, $t(281) = 3.90$, $p < .001$, and 120° *vs.* 180°, $t(281) = 6.85$, $p < .001$, but for level 2 judgements all angle contrasts were significant (all *t*s > 10.2, *p*s < .001), showing that the increasing angle of rotation has a stronger effect on response times for level 2 judgements. In addition, the Angle x Content interaction was significant. Follow-up analyses tested the effect of angle separately for each content type, and showed a significant effect of Angle on both spatial perspective-taking, $F(3, 843) = 160.92$, $p < .001$, $\eta_p^2 = .36$, and visual perspective-taking, $F(3, 843)= 269.65$, $p < .001$, $\eta_p^2 = .51$. Response times increased as the angle of rotation increased (all contrasts $p < .01$), though this increase was steeper for visual than spatial judgements.

Importantly, Age group modulated some of the effects reported above. The 2-way interactions between Age group x Angle and Age Group x Level were significant, and were further subsumed under a 3-way interaction between Age group, Angle and Level. Follow-up analyses showed that the Age Group x Angle interaction was significant only for level 2 judgements, $F(6,846) = 7.45$, $p < .001$, $\eta_p^2 = .05$, and not for level 1, $F(6,837) = 1.95$, $p = .07$. Next, we tested the effect of angle separately for each age group at level 2, revealing a significant effect of Angle in all three age groups for level 2 judgements [adolescent, $F(3, 267) = 179.18$, $p < .001$, $\eta_p^2 = .67$; young, $F(3, 309) = 268.51$, $p < .001$, $\eta_p^2 = .72$; old, $F(3, 261) = 132.72$, $p < .001$, $\eta_p^2 = .60$], with young adults showing the greatest effect of angle. Further follow-up analyses showed that the Age group x Level interaction was significant at angles 120°, $F(2, 279) = 6.52$, $p = .002$, $\eta_p^2 = .05$, and 180°, $F(2, 279) = 7.05$, $p = .001$, $\eta_p^2 = .05$, but not 0° or 60°, Fs < 2.7, ps > .07. Next, we tested the effect of level separately for each age group and angle, and showed that all participants had slower reaction times for level 2 than level 1 judgements at angle 120° [adolescent, $t(89) = 16.05$, $p < .001$; young $t(103) = 18.71$, $p < .001$; old $t(87) = 14.48$, $p < .001$] and 180° [adolescent, $t(89) = 19.60$, $p < .001$; young $t(103) = 21.03$, $p < .001$; old $t(87) = 16.07$, $p < .001$], but the effect of level at these angles was slightly smaller in the older adults than both the adolescent and young adult groups.


Level
1
2
Adolescents Spatial Content
Adolescents Visual Content
logRT(ms)
RT(ms)
3.2
3.1
3.0
2.9
2.8
0°
60°
120°
180°


Level
1
2
Young Spatial Content
Young Visual Content
logRT(ms)
logRT(ms)
3.2
3.1
3.0
2.9
2.8
0°
60°
120°
180°

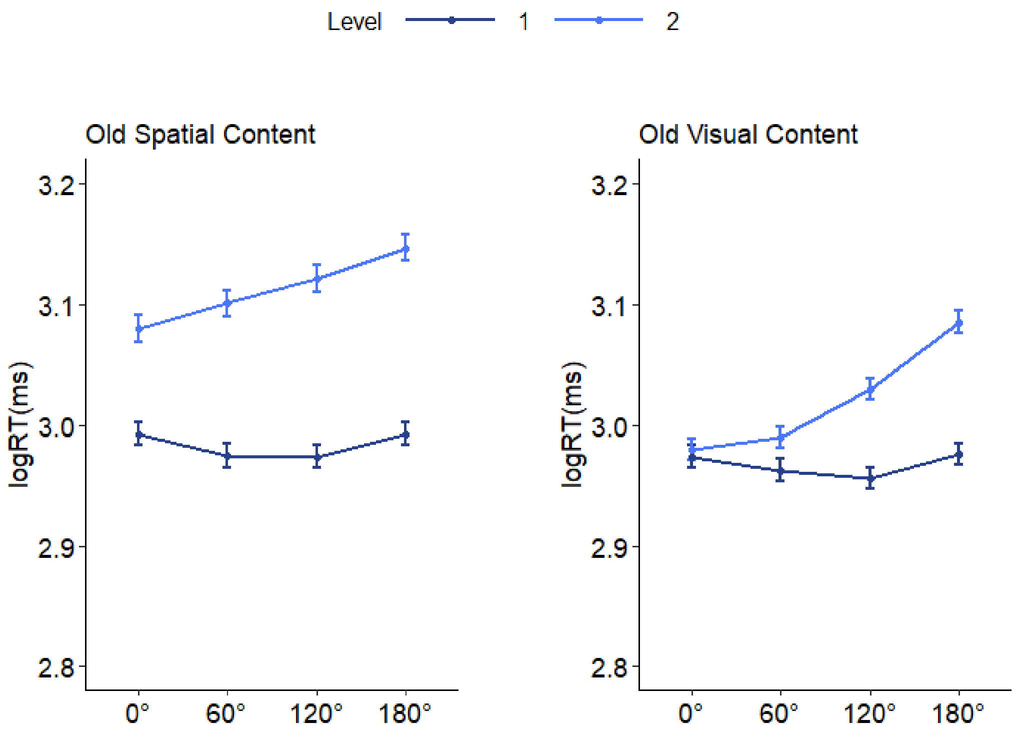


Figure 2.9: The log reaction time in each condition and age group in the visuo-spatial perspective-taking task, showing the condition mean and standard errors.

2.3.3. Interactive reference assignment

Only experimental trials were included in our analyses. Accuracy was calculated as the proportion of trials on which participants correctly selected the target object, and reaction times were calculated only on correct trials. Outliers were excluded from the analysis of reaction times if they fell more than 2.5 standard deviations from the age group's mean response time, or were faster than 200ms (since this indicated they selected the object before hearing the scalar contrast term). The final sample in this task was 281 participants (90 adolescents, 104 young adults, 90 older adults); one participant from the adolescent group was excluded from the analysis as the accuracy was lower than 50% and two participants due to computer failure. A 3 x 2 mixed design ANOVA was used for analysis of accuracy and log-transformed reaction time data, crossing the between-subjects factor Age Group (adolescents *vs.* young adults *vs.* older adults) with the within-subjects factor Condition (listener only *vs.* shared view). Full statistical effects are reported in Table 2. 8 for accuracy, and in Table 2. 9 for reaction times.

Table 2. 8: Statistical effects for Accuracy in the Director task. Asterisks show significance of effects, where *** $p < .001$.

	df	*F*	*P*	η_p^2
Age Group	(2, 278)	9.74	< .001***	.07
Condition	(1, 278)	32.54	< .001***	.11
Group x Condition	(2, 278)	7.95	< .001***	.05

Analysis of accuracy revealed significant main effects for both variables. The main effect of Age group showed that overall, young adults (*M* = .98) were more accurate at selecting the target object compared to older adults (*M* = .93; *t*(191) = 4.13, $p < .001$), and compared to adolescents (*M* = .96, *t*(190) = 2.41, *p* = .02). The main effect of Condition confirmed our prediction that participants would commit more errors in the listener only condition (*M* = .94[1]) compared to the shared view condition (*M* = .98; *t*(280) = 5.39, $p < .001$).

The Age group x Condition interaction was also significant. Follow-up analyses used 1-ANOVAs to test whether accuracy differed between the three age groups, separately for each condition. Results revealed that Age group modulated accuracy in the listener only condition, *F*(2, 278) = 9.47, $p < .001$, $\eta_p^2 = .06$, but not in the shared view condition, *F* (2, 278) = 1.91, *p* = .15. Planned comparisons between age groups in the listener only condition revealed that young adults were significantly more accurate than both older adults, *t*(191) = 3.94, $p < .001$, and adolescents, *t*(190) = 2.06, *p* = .04.

---

[1] Note that .056 of responses [adolescents *M*=.039, young *M*=.025, old *M*=.110] reflected egocentric errors (i.e. selecting the competitor object in the occluded slot), and .009 [adolescents *M*=.015, young *M*=.005, old *M*=.008] reflected other errors (i.e. selecting any other object, most likely due to inattention).

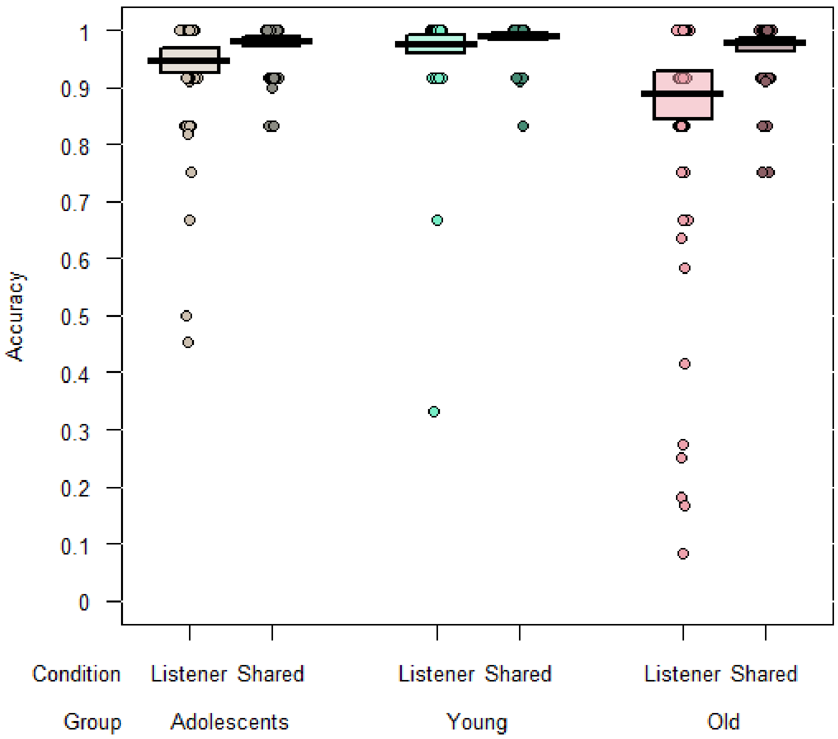


Figure 2.10: The accuracy in each condition and age group in the Director task. The plots show raw data points, a horizontal line reflecting the condition mean, and a rectangle representing the Bayesian highest density interval.

Table 2.9: Statistical effects for reaction times in the Director task. Asterisks show significance of effects, where *** *p* < .001.

	df	*F*	*P*	η_p^2
Age Group	(2, 278)	66.84	< .001***	.33
Condition	(1, 278)	1.92	.17	< .01
Group x Condition	(2, 278)	1.82	.16	.01

Analysis of reaction times revealed significant a significant main effect only for Age group, confirming our hypothesis that older adults would be slower to respond (*M* = 3616ms) compared to young adults (*M* = 2720ms; *t*(191) = 9.50, *p* < .001). Response times in younger adults and adolescents did not differ (*M* = 2671ms, *t*(190) = .66, *p* = .51).

Although the Age group x Condition interaction was not significant, we conducted exploratory analyses to test our hypothesis that older adults would show a greater condition effect compared to adolescents and young adults. Paired t-tests were used to compare reaction times between conditions, separately for each age group, and revealed that the older adults were significantly slower to respond in the listener only condition compared to the shared view condition, *t*(88) = 2.04, *p* = .04, but this condition effect was not significant for either young adults, *t*(103) = .07, *p* = .95, or adolescents, *t*(87) = .08, *p* = .94.

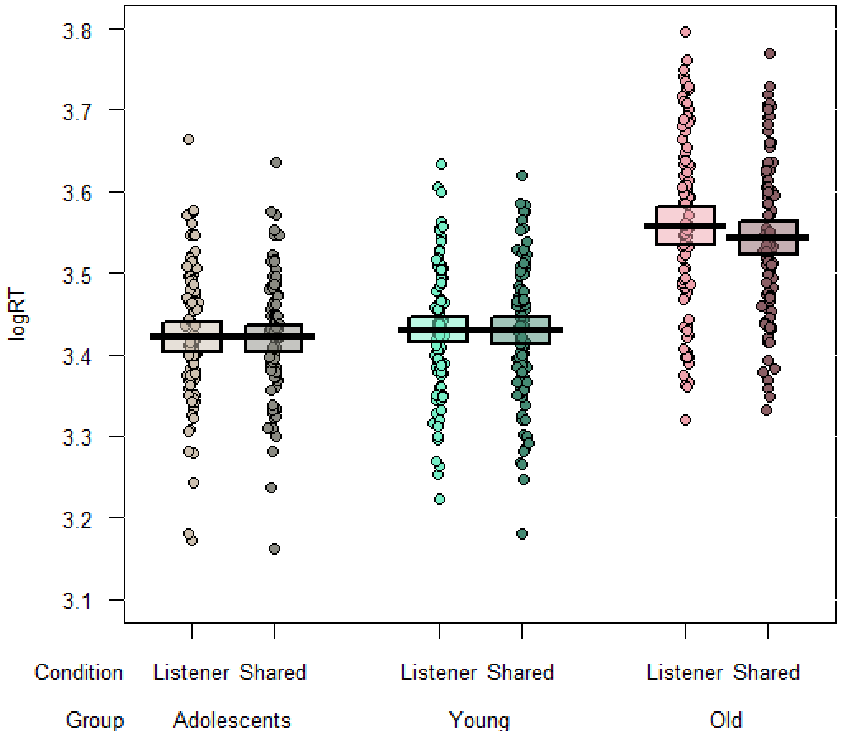


Figure 2.11: The log reaction time in each condition and age group in the Director task. The plots show raw data points, a horizontal line reflecting the condition mean, and a rectangle representing the Bayesian highest density interval.

2.3.4. Empathy for physical and social pain

The final sample in this task was 256 (76 adolescents, 94 young adults, 86 older adults); 16 participants were excluded due to too few segments (more than 25%), and seven participants were excluded due to excessive noise on the EEG recordings.

From pain ratings analysis, further two participants from the adolescent group were excluded due to computer failure in recording their ratings.

2.3.4.1. Pain ratings

Pain ratings were analysed using a 3 x 2 x 2 mixed design ANOVA, crossing the between-subjects variable Age group (adolescents *vs.* young adults *vs.* older adults) with the within-subjects variables Type (pain *vs.* no pain) and Content (physical *vs.* social). Full statistical effects are reported for pain ratings in Table 2.10.

Table 2.10: Statistical effects for pain ratings in the empathy for others' pain task. Asterisks show significance of effects, where * $p < .05$; ** $p < .01$; *** $p < .001$.

	Df	*F*	*P*	η_p^2
Age Group	(2, 251)	.76	.47	< .01
Type	(1, 251)	1092.82	< .001***	.81
Content	(1, 251)	44.94	< .001***	.15
Age Group x Type	(2, 251)	3.60	.03*	.04
Age Group x Content	(2, 251)	4.61	.01**	.03
Type x Content	(1, 251)	26.34	< .001***	.10
Age Group x Type x Content	(2, 251)	.76	.47	< .01

Results revealed significant main effects for the Type and Content of pain. The main effect of Type revealed that participants judged images depicting pain as more painful (M = 57.4) than no-pain images (M = 14.8, $t(253) = 33.13$, $p < .001$). The main effect of Content showed that participants judged physical stimuli as more painful (M = 39.0) than social stimuli (M = 33.2, $t(253) = 6.43$, $p < .001$). As expected, the Type x Content interaction was significant. Follow up t-tests compared the difference in pain ratings for pain and no-pain conditions (pain *minus* no-pain) between physical and social stimuli, and showed that the pain effect was larger when participants rated physical stimuli (M_{Diff} = 45.6) compared to social stimuli (M_{Diff} = 39.6, $t(246) = 5.05$, $p < .001$).

The 3-way interaction between Age group, Type and Content was not significant, however Age group modulated the effects of Type and Content separately. To follow up the significant Age Group x Type interaction, we used t-tests to compare the difference in pain ratings for pain and no-pain conditions (pain *minus* no-pain) between the three age groups. Results revealed that the type effect was larger in young adults (M_{Diff} = 46.73) compared to adolescents (M_{Diff} = 38.44, $t(166) = 2.71$, $p = .007$), but did not differ when comparing young adults and older adults (M_{Diff} = 41.65, $t(178) = 1.72$, $p = .09$). To follow up the significant Age group x Content interaction, we used t-tests to compare the difference in pain ratings for physical and social conditions (physical *minus* social) between the three age groups. Results revealed that the content effect was smaller in young adults (M_{Diff} = 2.23) compared to both adolescents (M_{Diff} = 7.46, $t(166) = 2.51$, $p = .01$) and older adults (M_{Diff} = 7.88, $t(178) = 2.73$, $p < .001$.

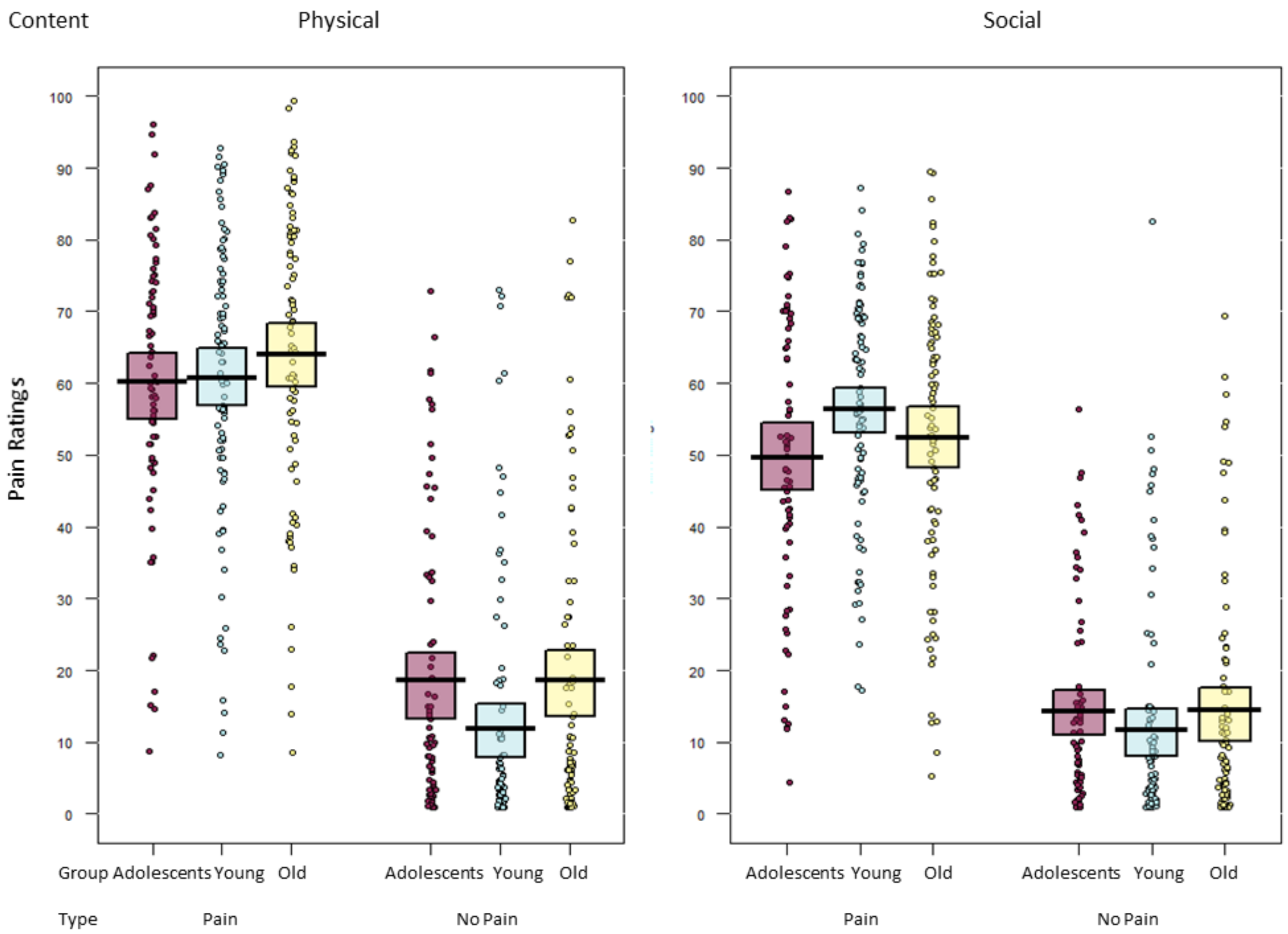


Figure 2. 12: The pain rating for each condition and age group in the empathy for others' pain task. The plots show raw data points, a horizontal line reflecting the content mean, and a rectangle representing the Bayesian highest density interval.

2.3.4.2 EEG analysis

Mu/alpha (8-13Hz) and beta (13-35Hz) suppression was analysed using separate mixed ANOVAs that crossed the between-subjects factor Age Group (adolescents *vs.* young adults *vs.* older adults) with the within-subjects factors Type (pain *vs.* no-pain), Content (physical *vs.* social) and Electrode site (central *vs.* occipital). Full statistical effects are reported in Table 2.11.

Table 2.11: Statistical effects for mu/alpha and beta suppression in the empathy for others' pain task. Asterisks show significance of effects, where * *p* < .05; ** *p* < .01; *** *p* < .001.

		df	*F*	*p*	η_p^2
Mu/Alpha suppression	Age Group	(2, 253)	7.74	< .001***	.06
	Type	(1, 253)	38.11	< .001***	.13
	Content	(1, 253)	17.33	< .001***	.06
	Electrode	(1, 253)	16.82	< .001***	.06
	Age Group x Type	(2, 253)	3.52	.03*	.03
	Age Group x Content	(2, 251)	.05	.95	< .01
	Age Group x Electrode	(2, 253)	21.11	< .001***	.14
	Type x Content	(1, 253)	.002	.96	< .01
	Type x Electrode	(1, 253)	2.90	.09	.01
	Content x Electrode	(1, 253)	.06	.81	< .01
	Age Group x Type x Content	(2, 253)	.76	.94	< .01
	Age Group x Type x Electrode	(2, 253)	2.37	.10	.02
	Age Group x Content x Electrode	(2, 253)	.43	.65	< .01
	Type x Content x Electrode	(1, 253)	.03	.86	< .01
	Age Group x Type x Content x Electrode	(2, 253)	1.26	.28	.01
Beta suppression	Age Group	(2, 253)	10.79	< .001***	.08
	Type	(1, 253)	12.58	< .001***	.05
	Content	(1, 253)	19.18	< .001***	.07
	Electrode	(1, 253)	6.11	.01**	.02
	Age Group x Type	(2, 253)	3.77	.02*	.03
	Age Group x Content	(2, 251)	1.69	.19	.01
	Age Group x Electrode	(2, 253)	1.83	.16	.01
	Type x Content	(1, 253)	2.05	.15	< .01
	Type x Electrode	(1, 253)	3.54	.06	.01
	Content x Electrode	(1, 253)	3.24	.07	.01
	Age Group x Type x Content	(2, 253)	2.56	.08	.02
	Age Group x Type x Electrode	(2, 253)	1.13	.33	< .01
	Age Group x Content x Electrode	(2, 253)	.41	.67	< .01

Type x Content x Electrode	(1, 253)	2.19	.14	< .01
Age Group x Type x Content x Electrode	(2, 253)	.74	.48	< .01

2.3.4.2.1. Mu/alpha suppression

Analysis of mu/alpha suppression revealed significant main effects for all variables. As expected, the main effect of Type revealed greater mu/alpha suppression for pictures that depicted pain (*M* = - .79) compared to no-pain (*M* = - .76). The main effect of Content showed that mu/alpha suppression was greater for physical stimuli (*M* = - .78) than social stimuli (*M* = -.76). The main effect of Electrode revealed that mu/alpha suppression was greater over occipital (*M* = -.79) compared to central electrode sites (*M* = -.75). The main effect of Age group revealed that young adults showed a greater mu/alpha suppression (*M* = -.81) compared to older adults (*M* = -.72, *t*(178) = 3.88, *p* < .001), but did not differ compared to adolescents (*M* = -.78, *t*(168) = 1.22, *p* = .22).

Crucially, Age group significantly modulated mu/alpha suppression in response to Type of pain. Follow-up analyses revealed that the effect of Type was only significant in the young adult, *t*(93) = 3.13, *p* = .002, and older adult groups, *t*(85) = 6.02, *p* < .001, but not in the adolescent group, *t*(75) = 1.90, *p* = .061. In addition, the Age group x Electrode interaction was significant. Follow-up analyses showed that mu/alpha suppression differed significantly between the three age groups over both central, *F*(2, 253) = 5.86, p = .003, η_p^2 = .04, and occipital electrode sites, *F*(2, 253) = 12.55, p < .001, η_p^2 = .09. Age contrasts showed that young adults showed greater mu/alpha suppression compared to older adults over occipital sites, *t*(178) = 4.58, *p* < .001, but the two groups did not differ over central sites, *t*(178) = 1.83, *p* = .07. In contrast, young adults showed greater mu/alpha suppression compared to adolescents over central sites, *t*(168) = 3.25, *p* = .001, but the two groups did not differ over occipital sites, *t*(168) = .02, *p* = .98.

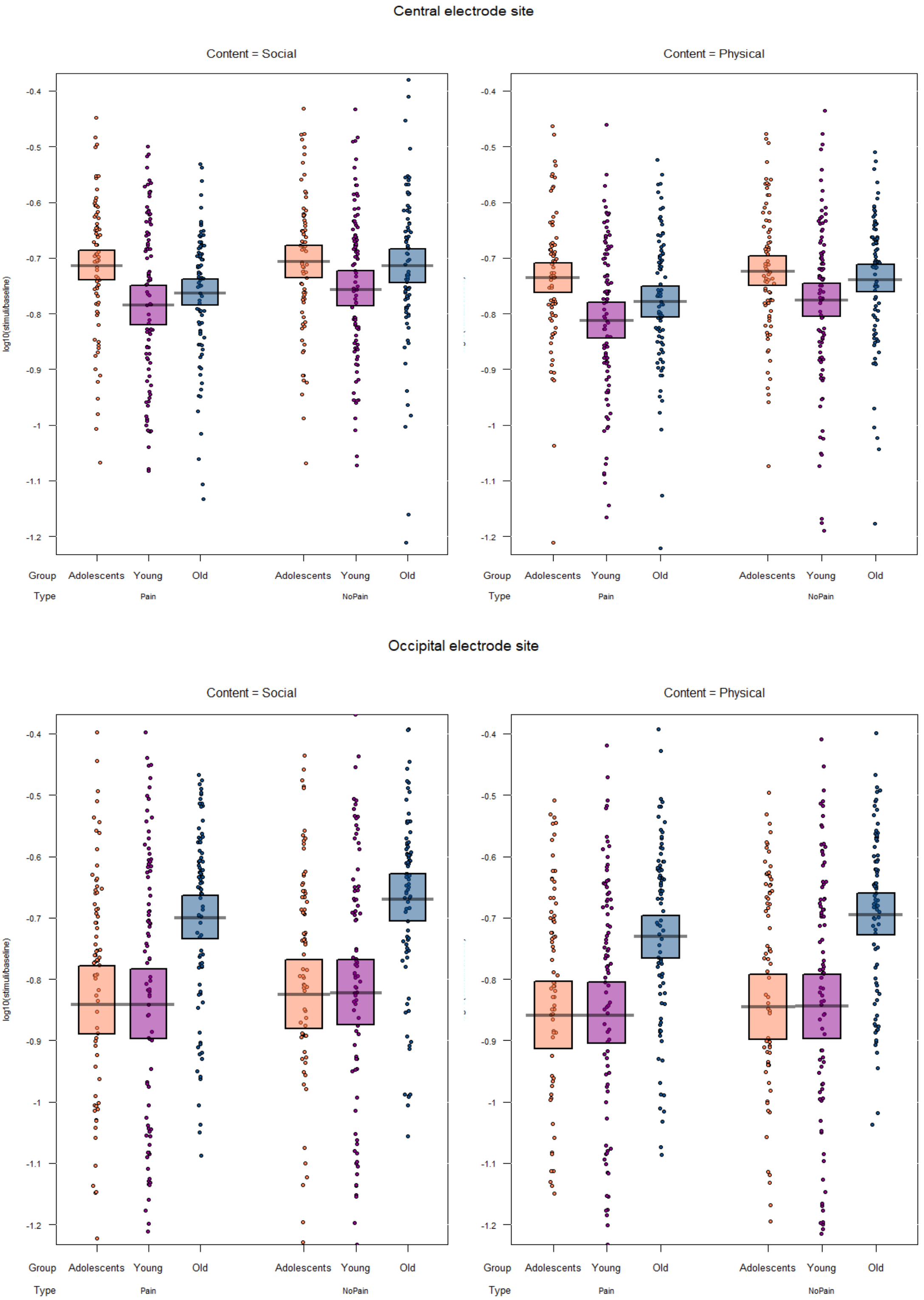


Figure 2.13: Mu/alpha suppression for each electrode site and condition in each age group in the empathy for others' pain task. The plots show raw data points, a horizontal line reflecting the content mean, and a rectangle representing the Bayesian highest density interval.

2.3.4.2.2 Beta suppression

Analysis of beta suppression revealed significant main effects for all variables. As expected, the main effect of Type revealed greater beta suppression for pictures that depicted pain (*M* = -.734) than no-pain (*M* = -.727 *t*(255) = 3.56, *p* < .001). The main effect of Content showed that beta suppression was greater for physical stimuli (*M* = -.74) than social stimuli (*M* = -.73, *t*(255) = 4.53, *p* < .001). The main effect of Electrode revealed that beta suppression was greater over occipital (*M* = -.74) compared to central electrodes (*M* = -.72, *t*(255) = 2.52, *p* = .01). The main effect of Age Group revealed that young adults showed a greater beta suppression (*M* = -.75) compared to older adults (*M* = -.70, *t*(178) = 4.19, *p* < .001), but did not differ compared to adolescents (*M* = -.74, *t*(168) = .31, *p* = .76).

Once again, Age group significantly modulated beta suppression in response to Type of pain. Follow-up analyses revealed that the effect of Type was only significant in the older adult group, *t*(85) = 5.25, *p* < .001, but not in the adolescent, *t*(75) = .63, *p* = .53, or young adult groups, *t*(93) = 1.15, *p* = .25.

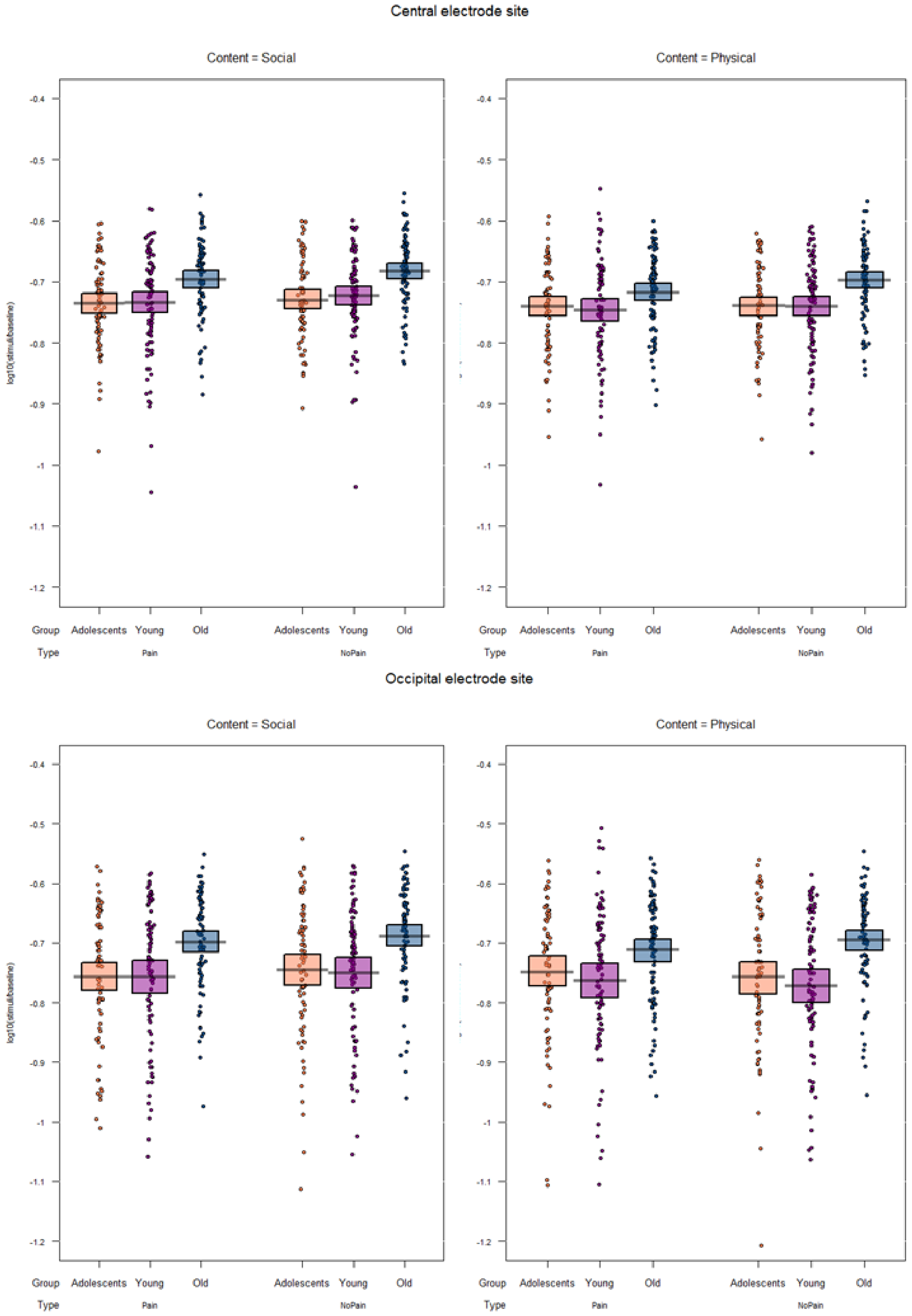


Figure 2.14: Beta suppression for each electrode and condition in each age group in the empathy for others' pain task. The plots show raw data points, a horizontal line reflecting the content mean, and a rectangle representing the Bayesian highest density interval.

2.3.5. Face to face conversation

2.3.5.1. Face to face conversation

The final sample in this task was 268 participants (89 adolescents, 99 young adults, 80 older adults); ten participants were excluded due to insufficient eye-tracking data and six participants were excluded due to technical issues. For the main analysis, fixations were analysed using a 3 x 2 x 3 mixed design ANOVA, crossing the between-subjects factor Age group (adolescents *vs.* young adults *vs.* older adults) with the within-subjects factors Condition (speaking *vs.* listening) and AoI (face *vs.* body *vs.* background). Full statistical effects are reported in Table 2. 12[2].

Table 2.12: Statistical effects for fixations in the face-to-face conversation task. Asterisks show significance of effects, where ** $p < .01$; *** $p < .001$.

	df	*F*	*p*	η_p^2
Age Group	(2, 265)	<.001	1	<.001
Condition	(1, 265)	<.001	1	<.001
AoI	(2, 530)	203.62	< .001***	.44
Age Group x Condition	(2, 265)	<.001	1	<.001
Age Group x AoI	(4, 530)	5.98	< .001***	.04
Condition x AoI	(2, 530)	128.71	< .001***	.33
Age Group x Condition x AoI	(4, 530)	3.18	.01*	.02

Results revealed a main effect of AoI, indicating that overall, participants distributed their attention differently towards the AoIs. Follow-up analyses showed that participants spent a greater proportion of time fixating the experimenter’s face

---

[2] Note that the main effects of Age Group and Condition, and the Age Group x Condition interaction are not meaningful in this analysis because proportions of fixations for each participant/condition summed to 1.

($M = .60$) compared to the background ($M = .24$; $t(267) = 13.87$, $p < .001$), and a greater proportion of time fixating the face features compared to the experimenter's body ($M = .15$; $t(267) = 17.55$, $p < .001$).

As expected, the Condition x AoI interaction was significant, showing that participants allocated their attention around the AoIs differently when speaking and listening. Follow-up analyses used t-tests to compare fixations for speaking and listening conditions separately for each AoI. As predicted, participants spent longer fixating the background while speaking ($M = .36$) compared to listening ($M = .13$, $t(267) = 15.40$, $p < .001$), but spent longer fixating the experimenter's face while listening ($M = .71$) comparing to speaking ($M = .50$, $t(267) = 11.44$, $p < .001$). In addition, participants spent longer fixating the experimenter's body while listening ($M = .17$) compared to speaking ($M = .14$, $t(267) = 2.25$, $p = .03$).

Importantly, the Age group x AoI interaction was significant, and was further subsumed under a 3-way interaction between Age group, Condition and AoI. Follow-up analyses showed that the Age group x AoI interaction was significant both while speaking, $F(4, 530) = 3.81$, $p = .005$, $\eta_p^2 = .03$, and listening, $F(4, 530) = 6.24$, $p < .001$, $\eta_p^2 = .05$. Therefore, we conducted separate 1-way ANOVAs for each AoI to compare fixations between age groups. Results for both listening and speaking conditions showed a significant difference between Age groups on fixations to the experimenter's face [listening: $F(2, 265) = 8.24$, $p < .001$, $\eta_p^2 = .06$; speaking: $F(2, 265) = 5.18$, $p = .006$, $\eta_p^2 = .04$] and the background [listening: $F(2, 265) = 9.06$, $p < .001$, $\eta_p^2 = .06$; speaking: $F(2, 265) = 4.52$, $p = .01$, $\eta_p^2 = .03$], but no effect of Age group on fixations to the experimenter's body, Fs $< .89$, ps $> .4$. Planned contrasts showed that young adults looked longer at the experimenter's face compared to both older adults [listening: $t(177) = 4.04$, $p < .001$; speaking: $t(177) =$

2.10, $p = .04$] and adolescents [listening: $t(186) = 2.69$, $p = .008$; speaking: $t(186) = 3.34$, $p < .001$]. Young adults subsequently spent less time looking at the background compared to both older adults [listening: $t(177) = 4.33$, $p < .001$; speaking: $t(177) = 1.95$, $p = .05$] and adolescents [listening: $t(186) = 2.79$, $p = .006$; speaking: $t(186) = 3.17$, $p = .002$].

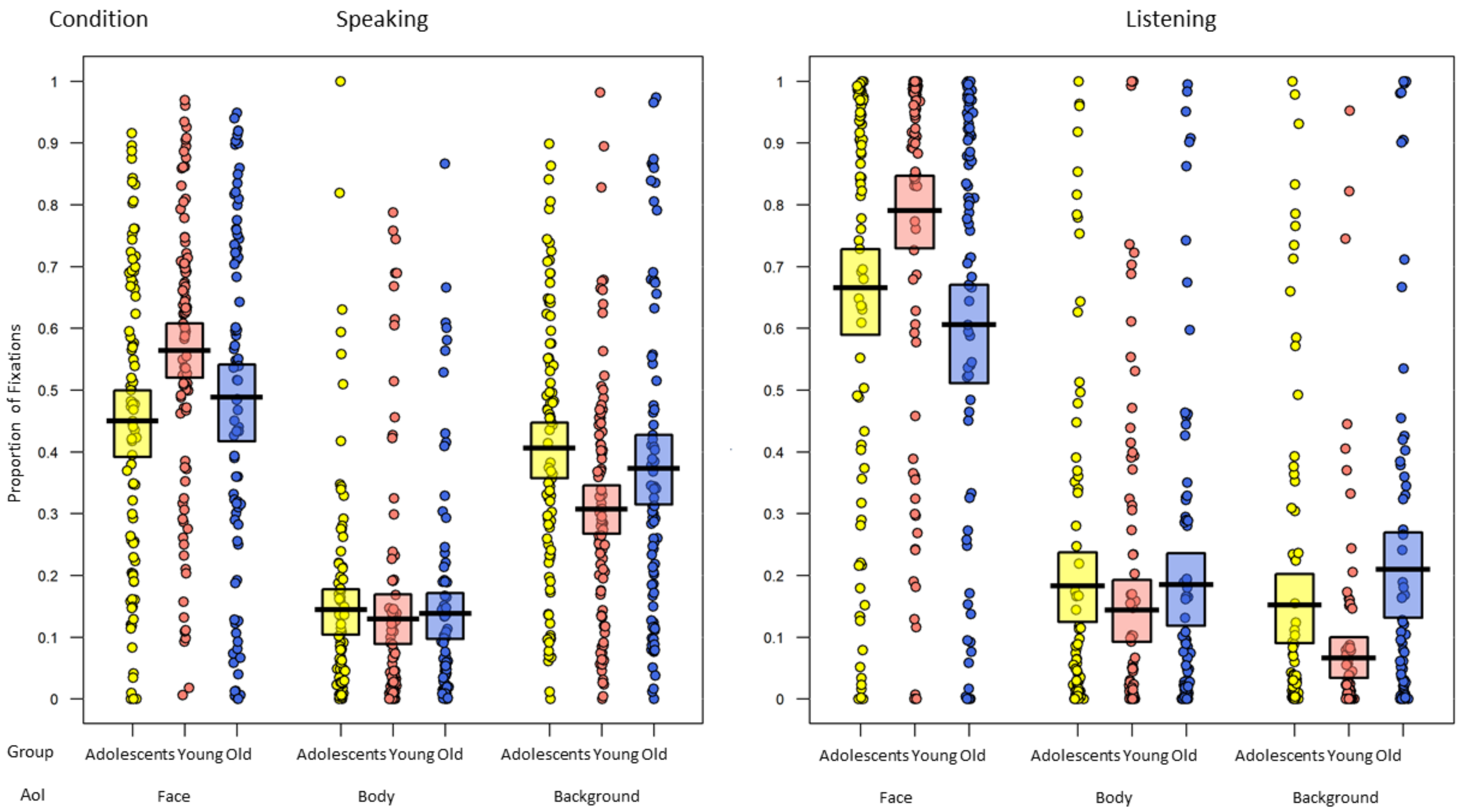


Figure 2.15: The proportion of time spent fixating each AoI in each condition and age group. The plots show raw data points, a horizontal line reflecting the condition mean, and a rectangle representing the Bayesian highest density interval.

2.3.5.2. Posters

Fixations to the three posters were analysed using a 3 x 2 x 3 mixed design ANOVA, crossing the between-subjects factor Age Group (adolescents *vs.* young adults *vs.* older adults) with the within-subjects factors Condition (speaking *vs.* listening) and AoI (shared gaze *vs.* averted gaze *vs.* neutral). Full statistical effects are reported in Table 2.13.

Table 2.13: Statistical effects for fixations to the posters in the face to face conversation task. Asterisks show significance of effects, where * $p < .05$; ** $p < .01$; *** $p < .001$.

	df	*F*	*p*	η_p^2
Age Group	(2, 265)	4.73	.01*	.03
Condition	(1, 265)	47.60	< .001***	.15
AoI	(2, 530)	6.28	.002**	.02
Age Group x Condition	(2, 265)	.52	.59	<.01
Age Group x AoI	(4, 530)	.98	.42	<.01
Condition x AoI	(2, 530)	.07	.93	<.01
Age Group x Condition x AoI	(4, 530)	1.88	.11	<.01

Results revealed a main effect of AoI, showing that overall, participants distributed their attention differently between the three posters. Follow-up analyses showed that participants spent a greater proportion of time fixating the neutral poster ($M = .03$) compared to both the averted gaze poster ($M = .02$; $t(267) = 2.59$, $p = .01$, and the shared gaze poster ($M = .02$; $t(267) = 2.80$, $p = .005$). Thus, contrary to our hypothesis, participants preferentially attended to posters that depicted non-social scenes compared to social scenes. The main effect of Condition confirmed that participants looked longer at the posters while speaking ($M = .03$) compared to listening ($M = .01$). Finally, the main effect of Age group was significant, showing that young adults spent less time looking at the posters ($M = .015$) compared to both older adults ($M = .024$; $t(177) = 2.19$, $p = .03$) and adolescents ($M = .03$; $t(186) = 3.11$, $p = .002$). None of the interactions were significant.

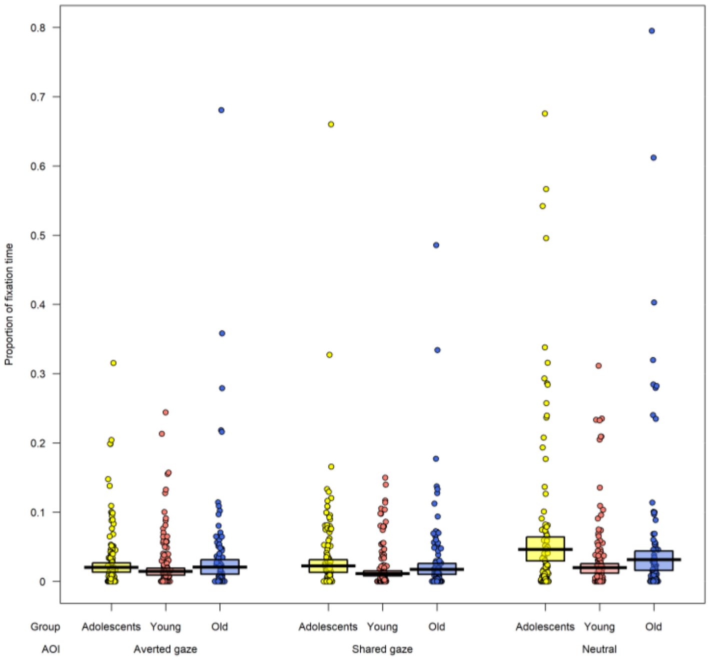


Figure 2.16: The proportion of time spent fixating each AoI in each age group. The plots show raw data points, a horizontal line reflecting the condition mean, and a rectangle representing the Bayesian highest density interval.

2.3.6 Navigating an environment

The final sample in this task was 271 participants (89 adolescents, 101 young adults, 81 older adults); six participants were excluded due to insufficient eye-tracking data and seven participants were excluded due to technical issues. Fixations were analysed using 3 x 4 mixed design ANOVA, crossing the between-subjects factor Age group (adolescents *vs.* young adults *vs.* older adults) with the within-subjects factor AoI (map *vs.* objects *vs.* path *vs.* people). Full statistical effects are reported in Table 2.14[3].

Table 2.14: Statistical effects for fixations in the navigation task. Asterisks show significance of effects, where * $p < .05$; *** $p < .001$.

	df	*F*	*p*	η_p^2
Age Group	(2, 268)	.01	1	<.01
AoI	(3, 804)	1494.41	< .001***	.85
Age Group x AoI	(6, 804)	3.01	.006**	.02

---

[3] Note that the main effect of Group is not meaningful in this analysis because proportions of fixations for each participant summed to 1.

Results revealed a main effect of AoI. Follow-up analyses compared the proportion of fixations on social (i.e. People) vs non-social stimuli (i.e. Path, Objects and Map). Participants looked less at people in their environment ($M = .05$) compared to any of the other AOIs: path ($M = .65$; $t(270) = 68.88$, $p < .001$), objects ($M = .11$; $t(270) = 12.15$, $p < .001$), map ($M = .19$; $t(270) = 15.61$, $p < .001$).

The interaction Age group x AoI was significant, and in line with our pre-registered analysis plan we conducted four 1-way ANOVAs, testing for differences between the three age groups, separately for each AoI. These analyses showed that fixations to people were modulated by Age group, $F(2, 268) = 5.67$, $p = .004$, $\eta_p^2 = .04$, as young adults spent more time looking at people ($M = .06$) compared to both older adults ($M = .04$, $t(180) = 2.80$, $p = .006$) and adolescents ($M = .04$, $t(188) = 2.85$, $p = .005$). Age group also modulated the time spent fixating the map, $F(2, 268) = 4.40$, $p = .01$, $\eta_p^2 = .03$, as young adults spent less time looking at the map ($M = .17$) compared to older adults ($M = .22$, $t(180) = 2.99$, $p = .003$), but did not differ compared to adolescents ($M = .19$, $t(188) = 1.53$, $p = .13$). The effect of Age group was not significant for any of the other AOIs (ps > .05).

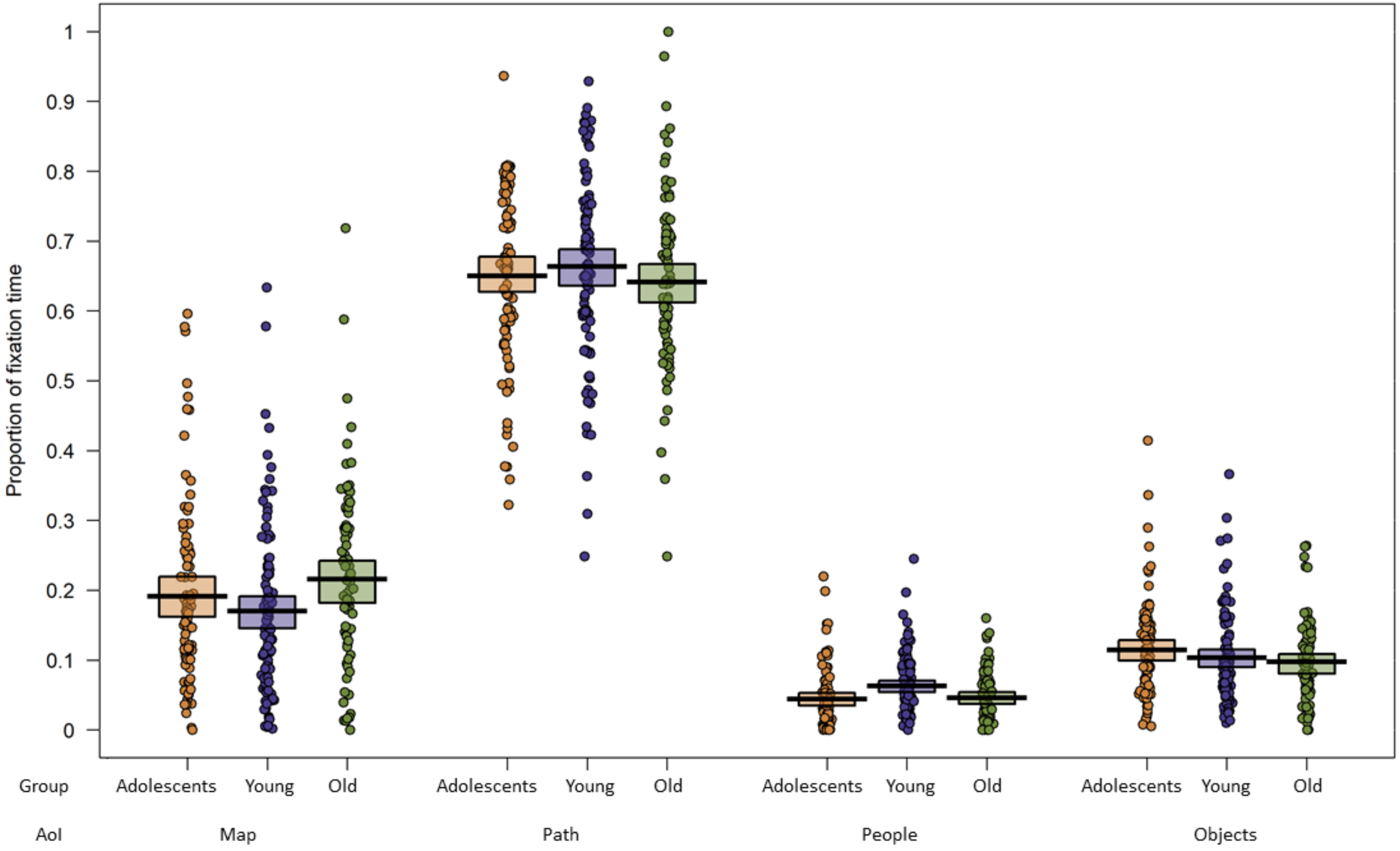


Figure 2.17: The proportion of time spent fixating each AoI in each age group. The plots show raw data points, a horizontal line reflecting the condition mean, and a rectangle representing the Bayesian highest density interval.

2.4. Discussion

All the basic effects of social cognition found in previous papers were replicated here, thus this discussion section will focus on age differences.

Overall, the results support the general hypothesis that social cognition abilities are enhanced among young adults compared to both adolescents and older adults, with four out of the six tasks eliciting some evidence for this claim. This provides strong evidence that age-related developments and declines in social cognition are not task-specific, and are therefore likely to represent broader changes in mentalizing ability. Interestingly, the pattern of results in the other two tasks suggested that older adults might be *more* sensitive to others' mental states compared to young adults and adolescents (i.e. higher likelihood of making social inferences and larger empathic brain responses to others' pain, see below), which reinforces the importance of assessing multiple sub-components of ToM and highlights differences in processing different domains. In addition, when response times were measured, older adults tended to be slower than either the young and adolescent groups, thus reflecting existing evidence for cognitive slowing in older age (Verhaeghen, 2011).

Results from the social inferences task showed that the likelihood and speed of making mental state inferences follows a comparable hierarchy across the lifespan (replicating Malle & Hollbrook, 2012). All age groups showed a relatively high likelihood of detecting mental states in the videos (~78% of the time), which demonstrates the human propensity to mentalize with other people. Inferences about intentions and desires were the most likely and faster to be detected, which suggests they might be activated automatically and reflect more basic aspects of social understanding. In contrast, understanding beliefs requires more effort, and all age groups were less likely and slower to make belief inferences, possibly because inferring someone else's beliefs involves an adjustment on one's own beliefs (Apperly, et al., 2008; Epley, et al., 2004a). Personality traits were the least likely and slowest to detect across all age groups. This is in line with previous studies that have shown personality inferences are harder to detect in lab settings compared to natural contexts, in which people are able to collect more information to draw personality traits (Buss & Craik, 1983; Wright & Mischel, 1987). In terms of aging, we found a linear effect whereby the likelihood of detecting social inferences increased from adolescence to young adulthood and older adulthood. These findings contrast with previous studies that have shown impaired mental state detection in older adults (Bailey & Henry, 2008; Henry, et al., 2013; Cavallini, et al., 2013), or that adolescents perform better than young adults (Fossati, et al., 2018). It is important to note, however, that in our study there were no correct or incorrect answers, and therefore our results might indicate the inability of older people to discriminate scenarios that involve mental states from those that do not require them, as found in Cavallini et al. (2013). Results from older adults might therefore reveal an over use of cause-effect inferences rather than enhanced social inference performance.

Results in the VSPT task provided some evidence for age-related changes in perspective-taking. Overall effects replicated previous work in showing a key role for the type of perspective-taking (i.e. level 1 *vs.* level 2) rather than the content (i.e. visual *vs.* spatial) that determined perspective-taking success; reaction times showed a steeper increase as angle increased for level 2 than level 1 judgements (Surtees et al., 2013a). Importantly, age modulated this interaction between level and angle.

Young adults showed steeper increments in reaction times with increasing angles when making level 2 judgements about how the avatar could see the cube/number; level 1 judgements did not differ with age. These results are in line with the proposal that level 2 judgements are more demanding of cognitive resources (Apperly & Butterfill, 2009), but also offer some evidence that not all perspective-taking is cognitively effortful and subject to an age-related development/decline in adolescence/older age. Specifically, level 1 judgements about *what* the avatar could see were unaffected by age, which suggests that they might be activated automatically.

Results from the Director task showed that participants were less accurate in the listener only than the shared view condition, confirming that participants of all ages experienced egocentric bias when taking another person's perspective (Epley et al., 2004a). Importantly, young adults outperformed both adolescents and older adults in the listener only condition, and only older adults showed a delay in responding in the listener only than shared view condition. This pattern supports our prediction for an extended period of social cognitive development in adolescence and a decline in older age (Blakemore & Choudhury 2006; Dumontheil et al., 2010; Mattan, et al., 2017). However, has previously mentioned compared to previous literature, our study has adopted a shorter version of this task with a reduced number of target trials thus this could have affected the performance.

Moreover previous studies have revealed a ceiling performance when analysing accuracy in this particular task (Ferguson & Cane 2017).

In the empathy for others' pain task, behavioural ratings revealed a greater pain effect (pain *vs.* no-pain images) among young and older adults compared to adolescents. In addition, physical pain was rated higher than social pain across all three age groups. Analysis of mu suppression in the EEG revealed greater alpha and beta suppression for pictures that depicted pain than no-pain, as expected (Chen, et al., 2012; Perry, et al., 2010; Yang, et al., 2009). More importantly, mu (alpha and beta) suppression in response to type of pain differed between the three age groups. In the alpha range, mu suppression to pain images was greater in the young and older adults compared to adolescents, and in the beta range mu suppression to pain images was greater among older adults compared to both adolescents and young adults. These findings are consistent with the proposal that empathy brain networks continue to develop through childhood (Decety, 2010) and adolescence (Levy & Feldman, 2017; Decety & Michalska, 2010), then remain stable or increase though adulthood and older age (Beadle & De La Vega, 2019). It is also important to consider that the age patterns seen here might reflect the use of adult actors in the images, which enhanced social closeness for our adult participants (Gutsell & Inzlicht, 2010). Finally, while global differences in mu suppression emerged between physical and social stimuli, this never interacted with the type of pain, which suggests that although behavioural responses distinguished different intensities for physical and social pain, neural activity did not.

In the face-to-face conversation task, we replicated the basic effects from Freeth et al. (2013) by showing that overall, participants preferentially attended to social content in the environment (i.e. the experimenter’s face), but this was modulated by phase of the conversation (i.e. more looks to the experimenter’s face while listening, and more looks to the background while speaking). More importantly, adolescents and older adults made fewer fixations towards their social partner compared to young adults, and in turn spent more time fixating the background compared to young adults. In addition, our incidental measure of social attention showed that young adults spent less time looking at the background posters compared to both adolescents and older adults, but that overall participants made more fixations to the poster depicting a non-social scene compared to either of the posters depicting social scenes (i.e. people with averted or shared gaze). In the navigation task we replicated the basic effects from Foulsham et al. (2011), showing that overall participants made more fixations on non-social content (i.e. objects, map, path) than social content (i.e. people) in the environment. More importantly, adolescents and older adults made fewer fixations towards people compared to young adults, and instead spent more time looking at the map. Thus, across both tasks we showed evidence that social attention is enhanced among young adults compared to both adolescents and older adults. The reduced social attention seen among adolescents and older adults in both tasks suggests that they experienced greater difficulty managing the cognitive effort of maintaining the conversation or route-finding compared to young adults (e.g. Barzy, et al., 2020; Doherty-Sneddon, et al., 2002; Doherty-Sneddon & Phelps, 2005; Glenberg, et al., 1998). Taken together, the finding that real-world social attention peaked in young adulthood is consistent with lab-based studies that have observed an extended period of development through adolescence and an older-age decline in ToM abilities (e.g. Blakemore, 2008; Bradford, et al., 2020; Brunsdon, et al., 2019; Henry, et al., 2013; Moran, 2013; Phillips, et al., 2002). Given that detecting social information in our environment and maintaining attention on it is an essential first step towards inferring other people’s mental states, we can conclude that social attention is a key mechanism to successful social interaction, and therefore diminished social attention is likely to be a primary source of the impaired ToM seen in adolescents and older adults compared to young adults.

2.5. Conclusions

To conclude, across six tasks that measured distinct sub-components of social cognition we found evidence that social cognition changes with advancing age. Interestingly, our results suggest a distinction between sub-components of social cognition that are cognitively demanding and therefore subject to a decline in older age (i.e. level 2 perspective-taking, reference assignment, social attention) and others that are relatively automatic and cognitively efficient (i.e. social inferencing, level 1 perspective-taking, empathy for others' pain) and are stable or even enhanced over the lifespan. The results therefore are in line with Apperly and Butterfill's (2009) two-systems account for ToM.

It has been hypothesised that age differences in ToM abilities are related to a general impairment in cognitive abilities (Henry, et al., 2013; Moran, 2013). In general, limited cognitive resources could explain that inhibiting the default self-perspective is more demanding, as well as distribute the attention on social stimuli. Overall performance at EFs seems to be a good predictor for successful ToM (Apperly, et al., 2009), therefore as mentioned in Chapter 1, previous investigations have been conducted to explore whether we can enhance social cognition through cognitive training.

With this in mind, we developed and validated a cognitive training protocol that was designed to enhance cognitive abilities but has the potential to improve social cognition. In the next Chapter I will present this adaptive cognitive training protocol, testing whether transfers within the cognitive domain are limited to the trained EF, as largely found from previous literature, or whether improvements generalize to other cognitive abilities.

CHAPTER 3: TRAINING EXECUTIVE FUNCTIONS USING A 21-DAY ADAPTIVE PROCEDURE AND ACTIVE CONTROL GROUP

This complete chapter has been submitted for publication:

De Lillo, M., Brunsdon, V., Bradford, E., Gasking, F., & Ferguson, H. J. (under review). Training executive functions using a 21-day adaptive procedure and active control group. *Quarterly Journal of Experimental Psychology*.

Training executive functions using a 21-day adaptive procedure and active control group

Martina De Lillo

Victoria E.A. Brunsdon

Elisabeth E.F. Bradford

Frank Gasking

Heather J. Ferguson

School of Psychology, University of Kent, U.K.

Acknowledgements

This work was carried out with the support of a European Research Council grant to HF (Ref: CogSoCoAGE; 636458). The pre-registration, datasets and code supporting this article are available on the Open Science Framework (https://osf.io/whxvt/).

3.1. Abstract

The degree to which executive function (EF) abilities (including working memory (WM), inhibitory control (IC), and cognitive flexibility (CF)) can be enhanced through training is an important question, however research in this area is inconsistent. Previous cognitive training studies largely agree that training leads to improvements in the trained task, but the generalizability of this improvement to other related tasks remains controversial. In this paper, we present a pre-registered experiment that used a ‘gold standard’ adaptive training procedure to examine whether EFs can be enhanced through cognitive training, and directly compared the efficacy and generalisability across sub-components of EF using training programs that target WM, IC or CF *vs.* an active control group. Participants (*n*=160) first completed a battery of tasks that assessed EFs, then were randomly assigned to one of four training groups, and completed a 21-day adaptive procedure that targeted a specific sub-component of EF (or was comparatively engaging and challenging, but did not train a specific EF). At post-test, participants returned to the lab to repeat the battery of EF tasks. Results revealed robust direct training effects (i.e. on trained task), but limited evidence to support near (i.e. same EF, different task) and far (i.e. different EF and task) transfer effects. Where indirect training benefits emerged, the effects were more readily attributable to the overlapping training/assessment task routines, rather than more general enhancements to the underlying cognitive processes or neural circuits.

Keywords: Executive functions; working memory; inhibitory control; cognitive flexibility; cognitive training; transfer effects

3.2. Introduction

Executive functions (EFs) is a commonly used ‘umbrella term’ to describe the set of processes that are responsible for higher-level action control (e.g. planning, inhibition, coordination and control of behaviours), and are necessary to maintain specific goals and resist distraction from alternatives. As such, EFs form the basis of our cognitive functioning (Diamond, 2013; Miyake, et al., 2000). In this paper we focus on three sub-components of executive functioning (inhibitory control, working memory, and cognitive flexibility) that contribute differentially to performance on complex executive tasks (Miyake, et al., 2000), and evaluate whether these EF abilities can be enhanced through a 21-day training procedure. Inhibitory control (IC) refers to the ability to inhibit a dominant response in order to focus on a more appropriate response; working memory (WM) is the ability to retain information for a brief period of time in order to perform mental operations; and cognitive flexibility (CF) is the ability to shift between different tasks or mental sets, and thus allows us to adapt behaviours to changes in the environment.

EFs have a protracted period of development, which begins in early childhood (~2 years old) and continues into young adulthood, with each sub-component of EF developing at its own rate (Diamond, 2002). For example, WM and planning have been shown to develop throughout childhood and into adolescence or early adulthood (e.g. Bishop, et al., 2001; Gathercole, et al., 2004), whereas CF and IC are thought to reach adult-like levels by age 12 (e.g. Crone, et al., 2006; Van den Wildenberg & Van der Molen, 2004). Neuroimaging findings support these distinct components of EF, and show that different goal-directed behaviours are subserved by distinct areas of the brain (e.g. Narayanan, et al., 2005; Crone, et al., 2006). Various clinical and neuro-degenerative conditions lead to impairments in EFs (e.g. Orellana & Slachevsky, 2013; Schmitt, et al., 2018; Stopford, et al., 2012), and even healthy ageing is associated with cognitive decline, which leads to general deficits in processing speed as well as specific impairments in EF, including WM (Braver & West, 2011), IC (Maylor, et al., 2011), and CF (Greenwood, 2007). Indeed, even when EFs are at their peak, there exists a great deal of individual variation in performance (Carroll & Maxwell, 1979). Understanding the conditions under which cognitive capacities function optimally, and how they relate to each other, is an important question.

Over the last decade or so, researchers have attempted to explore the degree to which EFs can be trained. Three main intervention approaches have been developed to improve cognitive functioning: i. strategy training, based on learning strategies to solve a task or function in everyday life (i.e. mnemonic strategies to enhance memory; Derwinger, et al., 2003); ii. multimodal training, considered as a more complex set of strategies that can, for example, involve physical activity and cognitive rehabilitation (Bherer, et al., 2013; see also Levine, et al., 2007); and iii. process training, which is centred on training a specific cognitive function (e.g. WM, processing speed). The cognitive training interventions tested here fall into this latter category.

The potential benefits of cognitive training have been linked to the concept of neuroplasticity. That is, through practice the brain reshapes its organization, creating new neural connections, or strengthening existing ones, which leads to reinforcement of the trained cognitive capacity and/or related abilities that were not trained directly (e.g. Ballesteros, et al., 2018; Han, et al., 2018; Jolles, et al., 2013; Kraft, 2012). Cognitive training is therefore based on the underlying principle that training on a specific task leads to improvements across the trained cognitive domain, and that this improvement might also extend to other related cognitive domains that were not trained. Three different classes of training effects have been identified (Carroll, 1993). *Direct* effects describe an improvement in performance on the trained measure, *near* transfer effects describe enhanced performance on a different measure of the trained construct (e.g. training and assessing WM using different tasks), and *far* transfer effects describe enhanced performance on a different construct (e.g. training WM and assessing IC). In order to obtain a transfer, it seems necessary that the trained task and the transfer task involve the same processes. Therefore, transfers are expected when there is overlap between the underlying processes involved in the different tasks (Dahlin, et al., 2008; Lustig, et al., 2009; Buschkuehl, et al., 2012).

To date, findings on whether EF abilities can be improved through training have been mixed (see Diamond & Ling, 2016, and Simons et al., 2016, for a review). On one hand, positive effects of direct training seem to be relatively uncontroversial, with research consistently showing that performance on a specific EF task improves with repeated practice (e.g. McKendrick, et al., 2014). On the other hand, the degree to which training effects transfer to untrained tasks or domains remains inconclusive. Some studies have reported near transfer effects within an EF domain (e.g. Karbach & Kray, 2009; Sandberg, et al., 2014; Heinzel et al., 2016), and others have shown far transfer effects between EF domains (e.g. Borella, et al., 2010; Dowsett & Livesey, 2000; Karbach & Kray, 2009; Salminen, et al., 2012), however many studies have reported no transfer effects at all (e.g. Blacker, et al., 2017; Holmes, et al., 2019; Melby-Lervag, et al., 2016; Owen, et al., 2010; Redick, et al., 2013). In fact, while one recent meta-analysis found small but significant improvements in fluid intelligence following WM training (Au, et al., 2015), two others have concluded that near transfer effects following training are weak and short-lived, and do not generalize across the sub-components of EF (Melby-Lervåg & Hulme, 2013; Simons, et al., 2016). Even more curiously, positive effects of cognitive training have been attributed by some researchers to placebo effects, where overtly advertising a study as examining the benefits of cognitive training biases participants' expectations and subsequent performance relative to participants who were recruited using non-suggestive advertising (Foroughi, et al., 2016). It is notable that the majority of research in this area so far has focused on outcomes following training in a single domain of EF, usually focusing on WM due to the rapid decline seen in this ability in older age (e.g. Heinzel, et al., 2014; Jaeggi, et al., 2010; Owen, et al., 2010; Redick, et al., 2013; Richmond, et al., 2011). Much less research has tested outcomes following training in other domains of EF, including IC and CF (Berkman, et al., 2014; Enge, et al., 2014; Karback & Kray, 2009; Thorell, et al., 2009), and studies very rarely compare training effects directly between different domains of EF.

One controversial aspect of research on cognitive training is the variability in methodology used across studies, which has limited definitive conclusions on the efficacy of EF training. These concerns have prompted leading researchers to make clear recommendations on the optimal approaches for EF training programs (Simons, et al., 2016; Diamond & Ling, 2016). First, it is important to assess baseline cognitive abilities before the training intervention (as well as after) so that the causal effects of training can be accurately quantified, and baseline differences between groups can be controlled. Second, training programs should include an active control group for comparison with the experimental group(s). Early studies in this area tended to use a passive control group, whose only contact with the experimenters was during the pre- and post- assessment sessions (e.g. Li, et al., 2010; Chein & Morrison, 2010; Salminen, et al., 2012), or engaged control participants in an activity that didn't match the cognitive demands of the experimental group (e.g. watching a film, playing videogames; Buschkuehl, et al., 2008; Borella, et al., 2010). Including an active control group allows researchers to rule out social or motivational factors that might elicit differences in performance between groups. Third, it is recommended to use adaptive training programs, where task difficulty increases as performance improves, since this challenges each participant to their own limits. Numerous studies have found superior training effects when task difficulty was adaptive *vs.* non-adaptive (e.g. Enge, et al., 2014; Lövdén, et al., 2010; Brehmer, et al., 2016; Klingberg, et al., 2002; Brehemer, et al., 2012; but also see Von Bastian & Eschen, 2016). Fourth, it seems clear that EF gains depend on the duration and frequency of the training; more training leads to better EF outcomes, but each training session should be relatively short (Au, et al., 2015; Jaeggi, et al., 2008). Lastly, studies should include large samples of participants who are randomly assigned to control/experimental groups, and have comparable expectations for improvement across groups.

In this paper, we present a pre-registered experiment that used a 'gold standard' 21-day adaptive training procedure to examine whether EFs can be enhanced through cognitive training, and for the first time directly compared the efficacy and generalisability across sub-components of EF of training programs that target WM, IC or CF *vs.* an active control group. Specifically, we compared performance on a battery of EF assessments before and after training to test for direct training effects (i.e. improvement on the trained task), near transfer effects (i.e. improvement on a different task that measures the same construct), and far transfer effects (i.e. improvement on a different task that measures a different construct). We specifically chose assessment tasks that differed in both paradigm and stimuli from the training task in each sub-component of EF, to ensure that any indirect training effects could not be attributed to shared strategies or response requirements between tasks. Training consisted of ten sessions over 21 days, completed at home through an online platform, with each training session lasting ~15 minutes (based on Enge, et al., 2014; Zinke, et al., 2014). Importantly, we used an active control group, in which participants completed a comparatively engaging and challenging task (an adaptive version of the lexical decision task) for the same duration as the EF training groups, and were blind to the different groups being tested. We tested a large sample (*n*=160 participants; 40 in each training group), and randomly assigned each participant to one of the four training groups.

Based on previous research summarised in Simons et al. (2016), we predicted that direct training effects would be apparent in all four training groups, i.e. performance on the trained task would improve from pre- to post-training. We also expected to observe small effects of near transfer in the three EF training groups, i.e. performance in the tasks that measured the same construct as the trained EF would improve from pre- to post-training. Finally, we tested whether training would lead to far transfer effects in the three EF training groups, i.e. performance in the tasks that measured a different cognitive construct to the trained EF would improve from pre- to post-training. We did not expect to find any far transfer training benefits in the control group, i.e. no improvement on any of the EF tasks from pre- to post-training.

3.3. Methods

All methodological procedures were pre-registered on the Open Science Framework (OSF) web pages (https://osf.io/whxvt/).

3.3.1. Participants

A total of 299 participants, aged between 18-35 years old, were recruited from the student population at the University of Kent, U.K. Of this total sample, 37 participants were excluded because they did not complete the online training sessions appropriately (i.e. under-training- less than nine sessions completed, or over-training- 12 or more sessions completed), 78 participants did not return to complete the post-training assessments, 16 participants were excluded as they were not native English speakers, and a further eight participants were excluded due to technical problems saving data. All participants were native English speakers, had normal or corrected-to-normal vision, had no known neurological disorders, and had no mental health or autism spectrum disorder diagnoses. Participants were randomly assigned to one of four training groups, with the final sample of n = 160 equally split between the four training groups (see Table 3.1 for demographic details per group), consistent with our pre-registered target sample size. The target sample of n = 160, 40 per group, was chosen a-priori based on similar research (e.g. Enge, et al., 2014; Zinke, et al., 2014), and a post-hoc power calculation showed that this sample yielded an estimated power of 87% with the significance level of α=.05 on 80% of occasions (as suggested by Cohen, 1988). Participants' consent was obtained according to the Declaration of Helsinki, and the Ethical Committee of the School of Psychology, University of Kent, approved the study.

Table 3.1: Participant demographics for each training group.

Training group	*n*	Mean age (*SD*)	M:F ratio
Inhibitory control	40	19.6 (3.9)	7:33
Working memory	40	20.2 (3.4)	7:33
Cognitive flexibility	40	19.2 (2.3)	3:37
Control group	40	19.3 (2.0)	5:35

3.3.2. Materials

3.3.2.1. Pre- and post-assessment tasks

All participants completed three assessment tasks in the lab during the pre- and post-training sessions.

Operation-Span (OSpan; Conway et al., 2005). This task was used to measure working memory (WM). Participants were asked to remember a sequence of letters that appeared one at a time on the computer screen (F, H, J, K, L, N, P, Q, R, S, T, and Y.). Between each letter, there was a distractor task (an arithmetical problem to solve). Participants were asked to recall the letters in the correct order at the end of each trial, clicking a box next to the appropriate letter(s) presented in a 4x3 matrix. A number appeared in the clicked box to indicate the order, and after completing the sequence of letters, participants received feedback on the correct number of the letters recalled. In cases where participants were not able to recall one or more letters, they were instructed to click on a blank box. Before the main task, participants familiarised themselves with the task through three practice blocks. The first block presented single letters in the middle of the screen for 800 ms, and participants had to memorize sequences of two or three letters. The second practice block required participants to solve some maths equations (e.g., (2 x 1) + 1 = 3), by indicating whether the answer was correct or incorrect as quickly and accurately as possible. In the last practice block, participants completed both the letter recall and maths tasks together. First, the maths equation was presented, followed by a letter appeared in the middle of the screen for 800ms; this sequence was repeated twice to create two-letter span trials, then the letter recall screen with the 4x3 letter matrix was presented. Participants completed three full practice trials, and were given feedback on how many letters they recalled correctly and how many errors they made on the maths problems. After completing the practice blocks, they started the experimental block which consisted of three trials for each of 2 to 7 letter spans (in a randomised order for each participant). This created a total of 18 trials with 81 maths problems and 81 letters. Participants were encouraged to keep their maths accuracy at or above 85% at all times. During recall, a percentage in red was presented in the upper right-hand corner of the screen, indicating the percentage accuracy for the maths problems. The dependent variable for this task was the *Partial Ospan Score*, calculated as the total number of letters correctly recalled, regardless of order.

Stroop Task (Stroop, 1935). This task was used as a measure of inhibitory control (IC). Participants were shown a series of words on the screen in one of four colours: RED, BLUE, GREEN and YELLOW. Colour words (e.g., BLUE) were presented in either a consistent or inconsistent colour (i.e., the word BLUE shown in blue/red ink, respectively). Neutral-words were also presented (e.g., CAT printed in green ink) to provide a baseline of colour naming without lexical interference. Participants were instructed to respond as quickly and accurately as possible to the ink colour of the words, ignoring the meaning, using a keyboard. Four coloured stickers indicated the four different colour responses: RED, BLUE, GREEN and YELLOW. Once the word appeared on the screen, participants gave their response and the next trial started immediately. After completing a practice block of 20 trials (10 neutral and 10 congruent), participants completed the experimental block which consisted of 50 congruent trials, 50 incongruent trials, and 50 neutral trials. Words were presented in a pseudo-randomised order, in which the same colour word, the same printed ink colour, or the same colour word/ink colour combination could not appear on two consecutive trials to avoid priming effects. We measured accuracy and reaction times for neutral, congruent and incongruent trials. For analysis, we calculated a *Stroop Effect* score for correct responses only; responses times were first transformed to z-scores, and the Stroop Effect was calculated by subtracting the mean RT for Congruent trials from the mean RT for Incongruent trials.

Wisconsin Card Sorting Task (WCST; Grant & Berg, 1948; Miyake et al., 2000). This task was used to measure cognitive flexibility (CF). Participants were asked to sort cards according to one of three classification rules: colour (red, blue, yellow, or green), shape (crosses, circles, triangles, or stars), or number of symbols (one, two, three, or four). A series of four cards appeared on the top of the screen which differed in colour, shape, or number of symbols, and one card appeared at the centre bottom. Participants had to figure out which of the three possible sorting rules to adopt according to the feedback that they received after choosing a card. Participants were told that the sorting rule would change throughout the task. There was no practice block, and the experimental block consisted of 128 cards. After clicking on a card, feedback was displayed on the screen stating whether the card had been sorted correctly or incorrectly. If incorrect feedback was received, participants had to switch to a different rule until they received correct feedback. After ten consecutive correct trials, the rule changed.

The dependent variable was the total number of perseverative errors, defined as the number of times in which participants persisted with an incorrect sorting rule.

3.3.2.2. Training tasks

Participants were randomly assigned to one of four groups, three that trained a specific component of EF (WM, IC, or CF), and an active control group (lexical decision task, LD). The training tasks were designed to be adaptive, in that task difficulty increased/decreased based on the participant's performance. Specifically, accuracy was monitored for each block so that if a participant's accuracy on that block equalled or exceeded 90% the task moved up to the next level of difficulty, and if accuracy equalled or fell below 75% the task returned to the previous level of difficulty. When accuracy on a block fell between 76% and 89% participants repeated the same level. Details of the levels of difficulty used for each of the training tasks are provided below. Participants received feedback on their accuracy at the end of each block. Practice blocks were excluded for training sessions completed at home. Training tasks were completed online. Each training session lasted approximately 15 minutes.

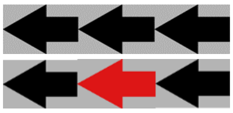
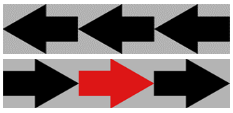
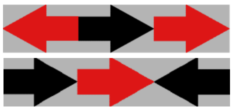
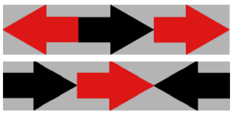
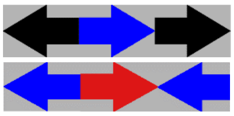
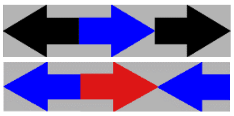
N-back task (Cohen, 1993). WM training adopted a visual version of the *n*-back task. A series of letters appeared one-by-one in the centre of the screen (500ms), and participants' task was to press a button on the keyboard if the letter presented was the same as the one presented *n* trials before. No response was required if the letter did not match. There were six different *n*-back levels: 1-, 2-, 3-, 4-, 5-, and 6-back (e.g., in the 2-back condition, participants should respond if the current letter is the same as the letter presented two trials before). Participants first completed three practice blocks with 1-, 2- and 3-back, then completed a further 15 blocks in the lab task or 21 blocks in the online task. Each block included 20 trials, with a fixed ratio of target/non-target trials of 6/14. Task difficulty increased over 15 levels by manipulating the *n*-back levels (between 2- and 6-back) and ISI (1800ms, 1600ms, and 1400ms), as shown in Table 3. 2. The dependent variables for this task were average level and accuracy, calculated as the proportion of Hits (i.e., correctly identifying a target as a target) minus False Alarms (i.e., incorrectly classifying a non-target as a target).

Table 3.2: Difficulty levels used in the *n*-back (WM) training task

Level	***n*-back**	**ISI (ms)**
1	2 back	1800
2	2 back	1600
3	2 back	1400
4	3 back	1800
5	3 back	1600
6	3 back	1400
7	4 back	1800
8	4 back	1600
9	4 back	1400
10	5 back	1800
11	5 back	1600
12	5 back	1400
13	6 back	1800
14	6 back	1600
15	6 back	1400

Stop Signal-Flanker task (Logan, 1994). IC training adapted the Stop Signal Task (SST) used in Berkman, Kahn, and Merchant (2014). Participants were presented with a black arrow in the centre of the computer screen, pointing left or right. Participants' task was to press the left or right arrow key on the keyboard to indicate the direction of this central arrow. However, on 25% of trials, the arrow turned red, after a variable stop signal delay (SSD). Participants were instructed to withhold their response on these trials. The task involved three short practice blocks, followed by nine blocks of experimental trials. Each block included 44 trials, with an equal ratio of left/right facing arrows. Task difficulty increased over eight levels by manipulating the presence and features of flanker stimuli and the length of the SSD (50-250ms, and 300-500ms), as shown in Table 3. 3. When flanker stimuli were present (levels 3-8), two additional arrows were placed either side of the central target arrow, and could either face the same direction as the central arrow (levels 3 and 4), or faced in a different direction to the central arrow (levels 5-8). On levels 3-6, these flanker stimuli could either be black or red, and did not change colour during the trial. Participants were instructed to ignore these distractor arrows and only respond to the direction and colour of the central arrow. An additional rule was added for the final two levels, as the arrows could appear in black, red or blue ink, and participants were instructed to withhold a response only for red colour changes to the central arrow (ignoring flanker arrows, and responding to blue colour changes to the central arrow). Thus, this task assessed two inhibitory control processes: the ability to withhold a response, and the ability to ignore competing stimuli (that may conflict with the target). The dependent variables for this task were average level and accuracy, calculated as the proportion of Hits (i.e., correctly identifying a black/blue arrow as a target) minus False Alarms (i.e., incorrectly classifying a red arrow as a target).

Table 3. 3: Difficulty levels used in the Stop Signal-Flanker (IC) training task.

Level	Trial Type	SSD (ms)	Example stimuli
1	Single arrow	50-250	
2	Single arrow	300-500	
3	Flanker arrows (same direction)	50-250	
4	Flanker arrows (same direction)	300-500	
5	Flanker arrows (different direction)	50-250	
6	Flanker arrows (different direction)	300-500	
7	Flanker arrows (different direction, additional colour)	50-250	
8	Flanker arrows (different direction, additional colour)	300-500	

Task Switching (Rogers & Mansell, 1995). Cognitive flexibility training adapted the task switching paradigm used in Barenberg, et al. (2015). Participants were presented with a 2x2 grid on the computer screen, and bivalent stimuli (a circle or triangle, in blue or yellow colour) appeared one-by-one in each of the four quadrants. Participants' task was to classify the stimuli by colour or shape, depending on trial type, using the keyboard. The task involved a short practice block, followed by 19 blocks of experimental trials. Each block included 32 trials, with an equal ratio of shape/colour combinations. Task difficulty increased over 12 levels by manipulating trial type (from single-task to mixed-task) and ISI (1250ms, 1000ms, and 800ms), as shown in Table 3. 4. In the single-task trial type, participants had to identify whether the stimuli colour was blue or yellow (levels 1-3), or whether the shape was a circle or triangle (levels 4-6). In the mixed-task trial type, participants indicated the stimuli's shape when it appeared in the upper two quadrants, and the stimuli's colour when it appeared in the lower two quadrants (thus had to switch categorization rule). Stimuli either appeared in a predictable clockwise manner (levels 1-9) or appeared in an unpredictable location in the grid (levels 10-12). The dependent variables for this task were average level and accuracy, calculated as the proportion of Hits (i.e. correctly identifying a target feature) minus False Alarms (i.e. incorrectly classifying a target feature).

Table 3. 4: Description of the levels in the Task Switching training protocol.

Level	**Trial type**	**Stimuli presentation**	**ISI (ms)**
1	Single-task (Colour)	Clockwise	1250
2	Single-task (Colour)	Clockwise	1000
3	Single-task (Colour)	Clockwise	800
4	Single-task (Shape)	Clockwise	1250
5	Single-task (Shape)	Clockwise	1000
6	Single-task (Shape)	Clockwise	800
7	Mixed-task	Clockwise	1250
8	Mixed-task	Clockwise	1000
9	Mixed-task	Clockwise	800
10	Mixed-task	Unpredictable	1250
11	Mixed-task	Unpredictable	1000
12	Mixed-task	Unpredictable	800

Lexical decision task (Meyer & Schvaneveldt, 1971). For the active control condition, we adopted a task that would be sufficiently cognitively taxing for participants, but wouldn't train any specific EF ability: the lexical decision task. In this task, participants used the keyboard to classify strings of letters as a word or non-word. A total of 3984 words were obtained using the MRC Psycholinguistics Database (Wilson, 1988), and were categorised according to their word frequency: High Frequency (HF), Middle High Frequency (MHF), Middle Low Frequency (MLF) and Low Frequency (LF). 7968 non-words were generated using the Wuggy pseudoword generator (Keuleers & Brysbaerd, 2010), retaining either one or two syllables from the matched real word (e.g. compare – cobbane – combore). The task involved a short practice block, followed by nine blocks of experimental trials. Each block included 40 trials, with an equal ratio of words and non-words, and each word was presented for 3000ms. Task difficulty increased over eight levels by manipulating word frequency (from high to low frequency) and the number of retained syllables for non-words (from one to two), as shown in Table 3. 5. The dependent variables for this task were average level and accuracy, calculated as the proportion of Hits (i.e. correctly identifying a word as a word) minus False Alarms (i.e. incorrectly classifying a non-word as a word).

Table 3. 5: Description of difficulty levels in the lexical decision training task, where HF = high frequency, MHF = middle high frequency, MLF = middle low frequency, LF = low frequency.

Level	Word frequency	Retained syllables	Example word / non-word
1	HF	1	Activity / Oupevici
2	HF	2	Activity / Aupetity
3	MHF	1	Compare / Cobbane
4	MHF	2	Compare / Combore
5	MLF	1	Expedient / Asquadent
6	MLF	2	Expedient / Ertopient
7	LF	1	Villainous / Nuttoilous
8	LF	2	Villainous / Nellailous

3.3.2.3. Motivation assessment

At the end of each online training session, participants completed a short questionnaire to assess their motivation to complete the task. This questionnaire was based on the Intrinsic Motivation Inventory (Deci & Ryan, 2015), and consisted of six statements (e.g. ‘I enjoyed doing this activity very much’, ‘I found this activity hard to complete’), which participants rated on a Likert scale from 0 (not at all true) to 7 (very true). Scoring was reverse coded where necessary to ensure that higher scores indicated greater motivation for each statement. An average motivation score, across all six statements and 10 online training sessions, was calculated for each participant.

3.3.3. Procedure

Participants first completed the 45-minute pre-training session in the lab, which included the three assessment tasks in a randomized order (i.e., OSpan, Stroop and WCST), followed by their assigned training task (either *n*-back, Stop-signal flanker, task switching, or lexical decision). At the end of the pre-training session, they received instructions on the procedures to complete the online training at home. Participants were invited to complete 10 online training sessions at home, each lasting ~15 minutes, over the next 21 days. From the final sample, 43 participants completed only nine training sessions at home (IC = 7; NB = 17; CF = 5; LD = 14), and five participants completed 11 training sessions at home (CF = 4; LD = 1). Training tasks were controlled through INQUISIT software (www.millisecond.com), and participants were sent personalized emails with a link to the appropriate task every two or three days. Following the 21-day training period[4], participants returned to the lab to complete the post-training session, in which they repeated the same three assessment tasks from the pre-training session, as well as their assigned training task.

3.4. Results

All analysis procedures were pre-registered, and the full datasets and analysis scripts are available on the Open Science Framework web pages (https://osf.io/whxvt/). All statistical analyses were conducted in R version 3.6.1. The dependent variables were z-scored for ease of comparison between tasks.

---

[4] We allowed a minimum of 20 days and a maximum of 28 days between pre- and post-training sessions (*M* = 21.4, *SD* = 1.1).

3.4.1. Direct training effects

To test our first hypothesis of performance improvements in all trained tasks when comparing pre- and post-training sessions, we conducted two mixed ANOVAs (one for accuracy and one for level), crossing the within-subjects variable Time (pre- *vs.* post-training) with the between-subjects variable Training group (WM *vs.* IC *vs.* CF *vs.* LD). Each dependent variable was z-scored over pre- and post-training, separately for each Training group. Data for accuracy and level are plotted, separately for each training group and pre-/post-training session, in Figure 3. 1. Data from each of the 12 training sessions in each training group is provided in the Appendix E for illustration.

3.4.1.1. Accuracy

Results revealed a significant main effect of Time, $F(1, 312) = 26.45$, $p < .001$, $\eta_p^2 = .09$, reflecting improved overall performance from pre- to post-training. Moreover, the interaction between Time and Training group was significant, $F(3, 312) = 8.04$, $p < .001$, $\eta_p^2 = .08$, suggesting that training effects differed between the four groups from pre- to post-training. To examine this interaction further, and following the pre-registered analysis, follow-up t-tests were conducted on pre-training *vs.* post-training outcomes separately for each group. Post-hoc tests showed that accuracy improved significantly from pre- to post-training in the WM, $t(39) = 7.16$, $p < .001$, and CF, $t(39) = 8.25$, $p < .001$, training groups, but did not improve significantly from pre- to post-training in the IC, $t(39) = .03$, $p = .97$, or LD control groups, $t(39) = .42$, $p = .67$.

3.4.1.2. Level

Results revealed a significant main effect of Time, $F(1, 312) = 120.72$, $p < .001$, $\eta_p^2 = .28$, reflecting improved overall performance from pre- to post-training. Moreover, the interaction between Time and Training group was significant, $F(3, 312) = 4.73$, $p = .003$, $\eta_p^2 = .04$, suggesting that training effects differed between the four groups from pre- to post-training. To examine this interaction further, and following the pre-registered analysis, follow-up t-tests were conducted on pre-training and post-training outcomes separately for each group. Post-hoc tests showed that the average of level difficulty improved significantly from pre- to post-training in all four groups [IC $t(39) = 4.73$, $p < .001$; WM $t(39) = 9.74$, $p < .001$; CF $t(39) = 13.75$, $p < .001$; LD $t(39) = 2.94$, $p = .005$], but was larger in the three EF training groups compared to the control group.

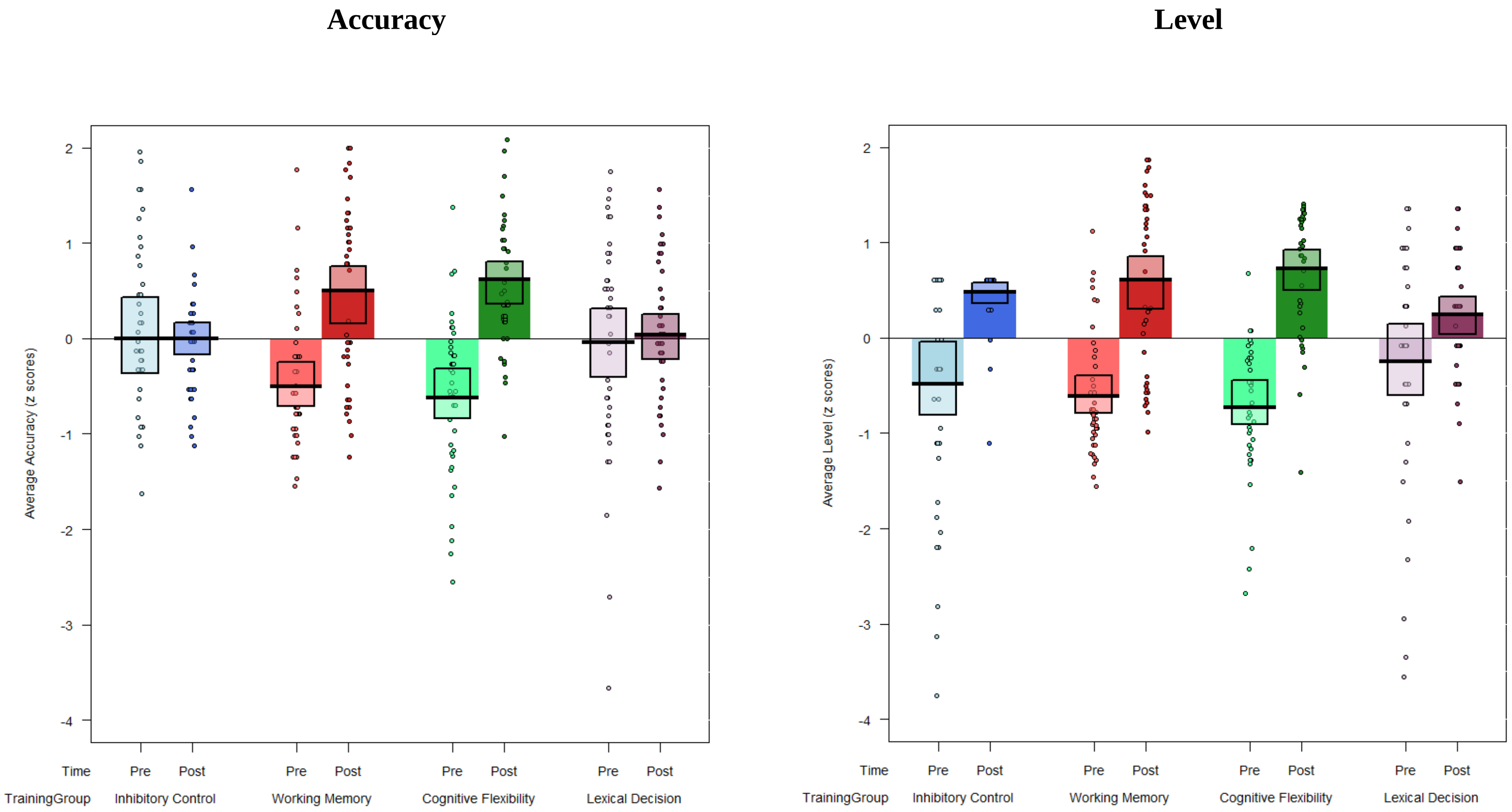


Figure 3. 1: Average z-scored accuracy (left panel) and level (right panel) pre- and post-training, plotted for each training group. Note that accuracy for each task is calculated as proportion of Hits *minus* False Alarms.

3.4.2. Indirect training effects

A series of ANOVAs were conducted on each of the assessment tasks to examine indirect training effects (i.e. near and far transfer), crossing the within-subjects variable Time (pre- *vs*. post-training) with the between-subjects variable Training group (WM *vs.* IC *vs.* CF *vs.* LD). Follow up t-tests were conducted to examine pre- and post-training performance in each of the assessment tasks separately for each training group. As per our hypotheses, these analyses examined near and far transfer effects of the trained cognitive ability. The perseverative errors for the WCST and the congruency effect for the Stroop task were reverse-scored so that a higher value indicates better performance for all assessment tasks. Data for each assessment task are plotted, separately for each training group and pre-/post-training session, in Figure 3. 2.

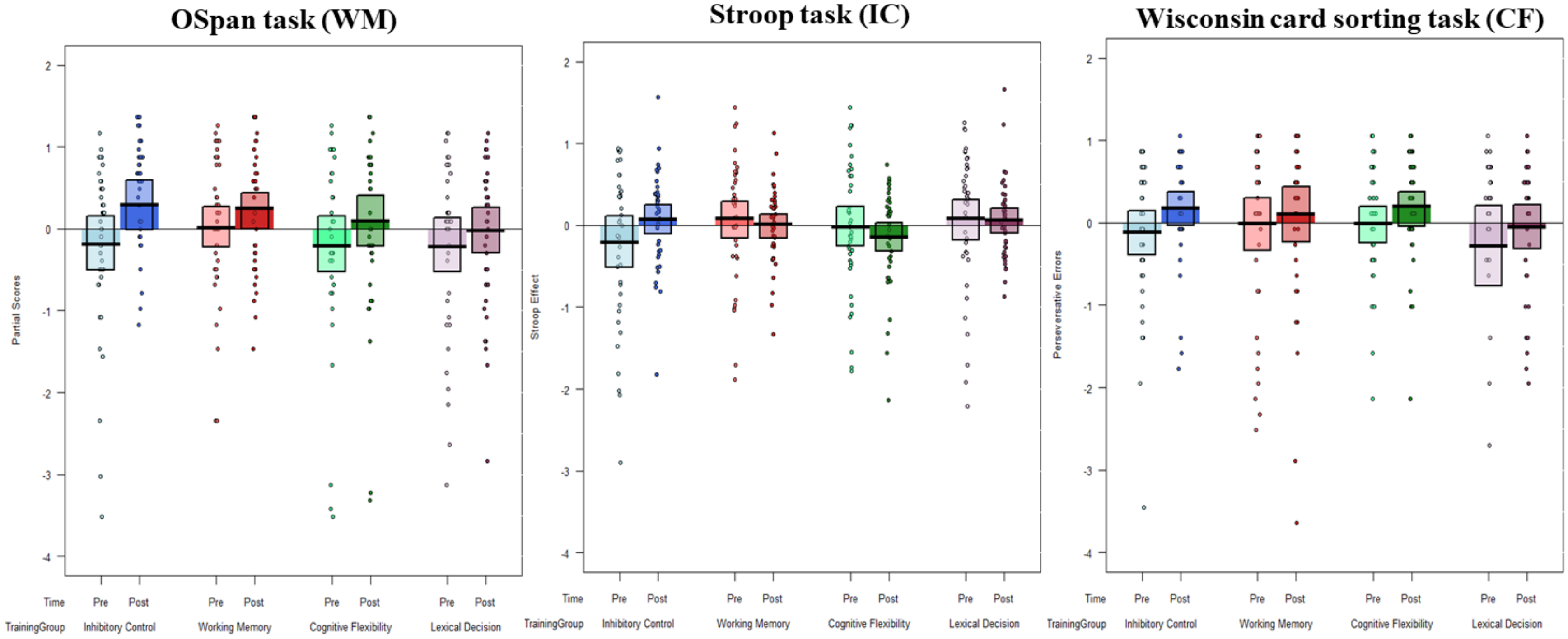


Figure 3. 2: Average z-scored partial scores on the OSpan task (left panel), Stroop effect on the Stroop task (middle panel), and perseverative errors on the WCST (right panel) at pre- and post-training, plotted for each training group.

3.4.2.1. OSpan / Working Memory

Results showed a significant effect of Time, $F(3,156) = 21.36$, $p < .001$, $\eta_p^2 = .12$, indicating that participants were more accurate (i.e. had higher partial scores) in the post-training session ($M = 68.6$) compared to the pre-training session ($M = 65.4$). The main effect of Training group was not significant, $F(3, 156) = .62$, $p = .60$, $\eta_p^2 = .005$, thus overall recall accuracy was comparable between the four training groups. The interaction between Time and Training group was also not significant, $F(3, 156) = .91$, $p = .442$, $\eta_p^2 = .017$. These results suggest that performance on the OSpan improved from pre- to post-training regardless of training group.

Following the pre-registered analysis, follow-up t-tests were conducted to compare pre- and post-training outcomes separately for each training group. For the WM training group, results showed a non-significant, improvement in partial scores from pre- to post-training, $t(39) = 1.90$, $p = .065$, indicating no significant near transfer from WM training to another measure of WM. However, OSpan partial scores significantly improved from pre- to post-training in both the IC, $t\ (39) = 3.83$, $p < .001$, and CF, $t\ (39) = 2.17$, $p = .030$, training groups. This suggests that training in IC and CF led to far transfer improvement on a measure of WM. Finally, there was no significant difference in OSpan partial scores from pre- to post-training in the LD training group, $t(39) = 1.44$, $p = .157$, indicating no improvement in WM in this active control group.

3.4.2.2. Stroop / Inhibitory control

Results revealed no significant effect of Time, $F(1, 156) = .058$, $p = .810$, $\eta_p^2 <.01$, indicating that the Stroop effect did not change from pre- to post-training. There was no significant main effect of Training group, $F(3,156) = .878$, $p = .454$, $\eta_p{}^2 =$

.02, and no significant interaction between Time and Training group, *F*(1,156) = 1.46, p = .227, η_p^2= .03. These results suggest that the Stroop effect did not change between pre- and post-training sessions, and did not differ between the four training groups.

Following the pre-registered analysis, follow-up t-tests were conducted to compare pre- and post-training outcomes separately for each training group. For the IC training group, results showed a non-significant, improvement in Stroop effects from pre- to post-training, $t(39) = 1.77$, $p = .085$, indicating no significant near transfer from IC training to another measure of IC. There were no significant differences in Stroop effects between pre- and post-training for either the CF, $t(39) = .89$, $p = .38$, or WM, $t(39) = .55$, $p = .585$, training groups, indicating no far transfer from CF or WM training to a measure of IC. Finally, there was no significant difference in the Stroop effect from pre- to post-training in the LD training group, $t(39) = .11$, $p = .915$, thus IC did not improve in this active control group.

3.4.2.3. WCST / Cognitive Flexibility

Results revealed a significant main effect of Time, $F(1, 156) = 5.45$, $p = .021$, η_p^2 = .034, indicating that participants made significantly fewer preservative errors post-training ($M = 5.04$) compared to pre-training ($M = 6.19$). The main effect of Training group was not significant, $F(3, 156) = 0.79$, $p = .503$, $\eta_p^2 = .029$, indicating that perseveration errors did not differ between the four training groups, and the interaction between Time and Training group was also not significant, $F(3, 156) = 0.17$, $p = .916$, $\eta_p^2 = .003$. Taken together, these results show that performance on the WCST improved from pre- to post-training regardless of training group.

Follow-up t-tests showed a non-significant, improvement in perseverative errors between pre- and post-training for the CF training group, $t(39) = 1.78$, $p = .082$, indicating no significant near transfer from CF training to another measure of CF. There was also no significant difference in perseverative errors between pre- and post-training for either the IC training group, $t(39) = 1.84$, $p = .073$, or the WM training group, $t(39) = .67$, $p = .504$, indicating no far transfer from IC or WM training to a measure of CF. Perseverative errors also did not differ from pre- to post-training in the LD training group, $t(39) = .92$, $p = .364$, indicating no improvement in this measure of CF in the active control group.

3.4.3. Motivation assessment

To check for group differences in motivation during the training, we conducted a one-way ANOVA on the questionnaire ratings (averaged over the six statements and 10 online sessions for each participant), with training group as the between-subjects factor. Note that data was missing for one participant (in the LD training group), thus analyses were conducted on a final sample of 159 participants. As can be seen in Figure 3. 3, results revealed a relatively moderate level of motivation among participants (overall $M = 3.8$), and importantly there was no difference in motivation between the training groups, $F(3,155) = .99$, $p = .41$, $\eta_p^2 = .02$ [IC $M = 3.73$, $SD = .85$; WM $M = 3.88$, $SD = .77$; CF $M = 3.93$, $SD = .98$; LD $M = 3.64$, $SD = .78$].

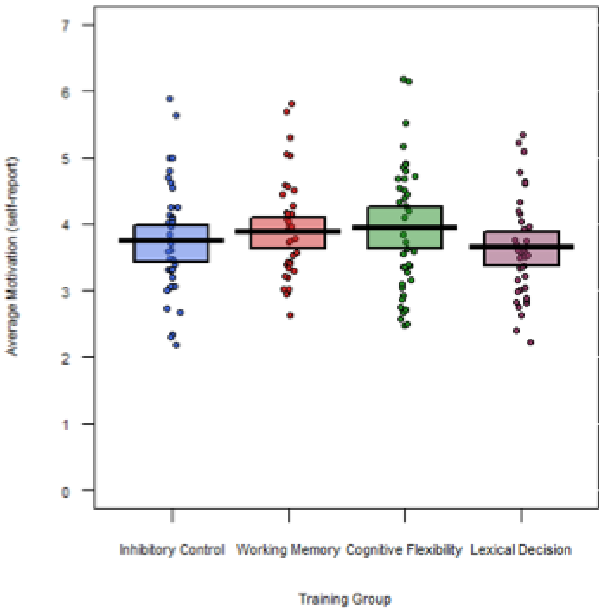


Figure 3. 3: Average self-reported motivation across the 10 online sessions in each training group.

3.5. Discussion

In this pre-registered experiment, we sought to examine whether EFs can be enhanced through cognitive training, and directly compared the efficacy and generalisability of training programs that targeted WM, IC or CF *vs.* an active control group. Participants (*n*=160) first completed a battery of tasks that assessed EFs, then were randomly assigned to one of four training groups, and completed a 21-day adaptive procedure that targeted a specific sub-component of EF (or was comparatively engaging and challenging, but did not train a specific EF). At post-test, participants returned to the lab to repeat the battery of EF tasks. Using this design, we were able to compare performance before and after training to examine direct training effects (i.e. improvement on the trained task), near transfer effects (i.e. improvement on a different task that measures the same construct), and far transfer effects (i.e. improvement on a different task that measures a different construct) in each training group.

In line with our predictions, all four training groups showed evidence of direct training; performance on the trained task improved from pre- to post-training. In the WM and CF training groups, this improvement was evident both in terms of their accuracy and difficulty level achieved, suggesting that repeated practice on the training task enhanced efficiency and ability in the trained EF measure. In the IC and active control groups, repeated practice on the SST/lexical decision task helped participants to achieve higher levels of difficulty (i.e. to ignore an increasing number of competing stimuli/distinguish lower frequency words and non-words with more retained syllables), but did not improve their overall accuracy. These effects are consistent with previous research, showing that practice improves performance on the trained task, or very similar tasks that involve the same strategies and response patterns (Noack, et al., 2009; Stine-Morrow & Basak, 2011).

In contrast, evidence for near transfer between different tasks that measured the same EF was very weak, and none of the effects reached statistical significance. These findings contrast with previous studies that have shown near transfer improvements following WM training (Heinzel, et al., 2014, 2016; Maraver, et al., 2016; Thorell, et al., 2009; Li, et al., 2010), IC training (Berkman, et al., 2014; Enge, et al., 2014; Thorell, et al., 2009), or CF training (Karbach & Kray, 2009). Importantly, however, our near transfer assessment tasks were specifically chosen to use different paradigms and stimuli to those used in the corresponding EF training task (i.e. WM was assessed with an OSpan task and trained with an *n*-back task; IC was assessed with a Stroop task and trained with a Stop Signal task; CF was assessed with the Wisconsin card sorting task and trained with a colour/shape switching task). This was done to isolate transfer effects on the cognitive process itself, and to avoid carry-over effects from shared strategies or response requirements between tasks. In fact, training benefits in previous research are strongest when the demands of the transfer task are highly similar to the trained task (e.g. Brehmer, et al., 2012; Harrison, et al., 2013; Soveri, et al., 2017), and effects are much less consistent when transfer tasks impose different processing demands (e.g. Blacker, et al., 2017; Gathercole, et al., 2019; Minear et al., 2016). These observations and the finding that near transfer effects were not evident in any of the three sub-components of EF tested here support the view that training effects seen in previous research most likely reflect specific features of the trained/assessment tasks and cognitive routines learnt during training, rather than more fundamental training benefits to the underlying cognitive ability (Gathercole, et al., 2019). The specificity of near transfer effects has been highlighted by a recent study that systematically manipulated the paradigm (*n*-back or complex span) and stimuli (verbal or visuo-spatial) used in a WM training program, and reported no transfer effects between paradigms, even when stimuli were matched (Holmes et al., 2019).

Similarly, evidence for far transfer between the different sub-components of EF tested here was weak. WM training did not lead to any improvements on tasks that measured IC or CF, nor did IC training or CF training alter performance on measures of CF or IC, respectively. However, our analysis revealed a significant improvement is the IC and CF groups, these need to be interpreted with caution due to the not significant interaction Training group and Time.

Taken together, this research contributes to theoretical debates on cognitive training and transfer effects. Transfer effects are thought to occur when the skills learnt in one domain generalise to enhance performance in another domain, and the degree to which this transfer occurs is directly related to the extent of shared features between the trained and untrained task (Singley & Anderson, 1989; Woodworth & Thorndike, 1901). This process is mediated by neural plasticity, as quantifiable changes emerge in the cortical and sub-cortical brain areas that subserve the trained cognitive ability through practice (Dahlin, et al., 2008). These shared features view would predict that near transfer effects are more likely and stronger than far transfer effects, since in the former the trained and untrained abilities share more common features (i.e. should rely on related cognitive and neural mechanisms). More recently, researchers have emphasised the role of the learning context, and its interaction with the content of the learned ability, suggesting that transfer effects across domains depend on the success of applying principles or strategies that are shared between the different tasks (Barnett & Ceci, 2002; Gathercole, et al., 2019). As such, transfer across cognitive domains relies on participants learning new skills that can be applied in similarly structured tasks. The finding in the current experiment that indirect training benefits were more likely to emerge following training in a *different* sub-component of EF (i.e. far transfer) than the *same* sub-component of EF (i.e. near transfer) goes against traditional shared features accounts of transfer, and instead suggests that transfer effects are mediated by cognitive routines learned during training and shared with the transfer task. Further research is needed to test how far training effects can transfer when shared cognitive routines are used to reinforce learning of the novel task.

Finally, we note some limitations with the current experiment, and propose some important avenues for future research in this area. First, some of the training tasks may not have been challenging enough for our highest performing participants, meaning that a ceiling level was reached (as seen in the SST training task accuracy), and the crucial adaptive aspect of the training was not consistent across participants/groups. Relatedly, the four training tasks differed in the number of adaptive levels included (ranging from eight to 15), and the degree to which increasing difficulty between levels was matched between training tasks, which may have limited comparability between training groups. The current tasks were selected based on those most commonly used in the field of cognitive training, and to avoid overlapping procedural elements between tasks. However, it is noted that some of the specific tasks used here are likely to have activated multiple sub-components of EF and are therefore limited in terms of cognitive specificity (as discussed for the OSpan task above). For instance, previous research has highlighted issues with using the Stroop task as a near-transfer measure of IC since it is more complex than most other measures of IC, requiring high levels of cognitive control to manage attention and semantic processing (Enge, et al., 2014). Building on the current research, future studies should aim to take a more systematic approach to controlling for similarities/differences in sub-routines between tasks to isolate key components that lead to training effects (as in Holmes et al., 2019). Second, our experiment tested a young adult student population, who are at their peak of cognitive functioning (Diamond, 2013; Hartshorne & Germine, 2015), and therefore the results may not be generalisable across the general population. Notably, previous research has observed larger training gains in groups whose cognitive abilities are not at their peak, for example among older adults (Karback & Verhaeghen, 2014), children (Zhao, et al., 2018), or clinical groups (e.g. Holmes, et al., 2010; Leśniak, et al., 2019; Hallock, et al., 2016). Finally, it has been suggested that to see widespread benefits of training on cognitive capacities, diverse skills must be practiced. Future studies could therefore adopt more varied training programs that tap multiple processes within a specific sub-component of EF (e.g. maintaining, updating and recalling in WM), while reducing the contributions of other EF sub-components, to allow a more rigorous exploration of ‘brain training’ and its generalisability across wider domains of functioning (e.g. social interaction; Kloo & Perner, 2003; Santiesteban, et al., 2012).

In sum, we conducted a pre-registered experiment that sought to adopt the ‘gold standard’ approach to EF training programs recommended by Simons et al. (2016), and investigated the efficacy and generalisability of EF training within and between three sub-components of EF (WM, IC and CF) compared to an active control training program. In line with previous literature we found robust direct training effects, but limited evidence to support near and far transfer effects (Heinzel, et al., 2014, 2016; Owen, et al., 2010). Where indirect training benefits emerged, the effects were more readily attributable to the overlapping training/assessment task routines, rather than more general enhancements to the underlying cognitive processes or neural circuits. Further research is needed to isolate sub-components of EF targeted in training programs, while systematically manipulating paradigm-specific commonalities between tasks. Such an approach would allow researchers to further explore what kinds of training most reliably lead to performance changes, and to assess the generalisability and specificity of training effects on cognition and beyond.

CHAPTER 4: TRAINING SOCIAL COGNITION INDIRECTLY THROUGH EXECUTIVE FUNCTIONS

4.1. Introduction

In earlier chapters I have shown that some key aspects of social cognition change with advancing age, and that there may be a cognitive basis for these changes. As discussed in Chapter 1, previous research has provided strong evidence for a relationship between social cognition and EFs, though the direction of this link remains under debate. Some of this work has suggested that preserved cognitive abilities in older age might mediate a substantial age-related decline in social interaction (e.g. Bailey & Henry, 2010; Cho & Cohen, 2019; Rakoczy, et al., 2012), while others have proposed that greater social integration and emotional support relate positively to cognitive function in older age (e.g. Béland, et al., 2005; Seeman, et al., 2001). Chapter 2 highlighted that level 2 perspective-taking, reference assignment, and social attention are impaired among adolescents and older adults compared to young adults, likely due to the cognitively demanding nature of these tasks. In Chapter 3, I tested the efficacy of cognitive interventions to enhance cognitive capacities in young adults, finding mixed evidence for training (i.e. robust evidence for direct training but only weak evidence for near and far transfer to other measures of EF). In this chapter, I further test the generalisability of these cognitive training protocols, examining whether training the relevant EFs can lead to indirect improvements in ToM.

For many years, researchers have attempted to improve social cognitive abilities through training in children (Lecce, et al., 2014; Slaughter & Gopnik, 1996), healthy adults (Trautwein, et al., 2020; Santiesteban, et al., 2012), clinical populations (Begeer, et al., 2011), and older adults (Cavallini, et al., 2015; Lecce, et al., 2015). This work has largely focused on interventions that aim to improve ToM directly (some of which was reviewed in Chapter 1), including engaging participants in conversation about mental states, mindfulness, perspective-taking, or role play. However, the questions of whether skills can generalise from a trained task to other tasks that measure related processes (including ToM and everyday mindreading) remains under debate. For example, some studies support improvements in social interaction abilities following engagement in a training protocol (e.g., Lecce, et al., 2014), whilst other studies show no significant improvements in social cognition performance following training interventions in children with Autism (Begeer, et al., 2011). A recent review highlighted the specificity of social training effects on behaviour, brain and physiology (Singer & Engert, 2019). In addition, factors such as age, motivation, verbal knowledge, and executive functions have been identified as important mediators to the success of training interventions (e.g. Lecce, et al., 2017; Zhang, et al., 2017).

Building on the developmental association between EF skills and ToM outlined in Chapter 1, a few researchers have begun to explore the effectiveness of training ToM indirectly by training the relevant EFs. Specifically, Kloo and Perner (2003) trained 3 and 4-year-old children on a set-shifting task (requiring cognitive flexibility) that required children to sort cards according colour and number dimensions. Results showed that training on the card sorting task significantly improved children’s performance on a test of false belief understanding. Similar effects have been demonstrated in a study on children with autism, where training on a set-shifting task led to a notable improvement in ToM (though ToM was not enhanced in their daily lives; Fisher & Happé, 2005). More recently, one study (Santiesteban, et al., 2012) has investigated whether these transferable training effects exist in adults (mean age 27 years). Here, participants completed a single session of training on one of three tasks: (1) imitate another person's hand action, (2) supress imitation of another person's hand action, or (3) 'general inhibition training' (i.e. performing hand actions based on colour cues). Results showed that while training social inhibition (i.e. group 2) enhanced participants' perspective-taking ability on a subsequent task, imitation and general inhibition training did not. However, further rigorous testing is required to elucidate the validity of these effects. Therefore, despite the potentially promising training opportunity that these studies suggest, research on the efficacy of EF training for improved ToM remains extremely limited, and no work to date has extended these findings using a broader range of EF training tasks, or across a wider age-range of individuals.

Studies on training typically suffer from a range of limitations, such as testing effects in very young children or clinical populations (which might limit the tasks and measures that can be used), healthy young adult participants (who may already be performing at or close to ceiling), focusing on just one aspect of social cognition, or adopting small sample sizes. All of these issues are likely to contribute to the mixed evidence on the efficacy of social cognitive training interventions, meaning that improvements are more difficult to detect. Indeed, larger training gains have been observed in groups whose cognitive abilities are not at their peak, such as among older adults (Karback & Verhaeghen, 2014) or children (Zhao, et al., 2018), and the nature of transfer effects to other domains can be distinct between different age groups (Heinzel, et al., 2014).

In this Chapter, we tested whether cognitive and social abilities can be enhanced by training the underlying EFs (using the protocols developed in Chapter 3 that targeted WM, IC and CF compared to an active control training group), and whether these training effects differ in different age groups (adolescents, younger adults and older adults). First, we expected to replicate the overall effects of age on social cognition from Chapter 2, distinguishing between those components of social cognition that are cognitively effortful and subject to age-related decline, and those which are relatively cognitively efficient and spared from age-related decline. Regarding cognitive abilities we expected to replicate the results reported in Chapter 3, showing limited effects of indirect transfer between tasks or EFs in the four training groups. More importantly, regarding social skills, we expected that training enhancement effects would be most likely in those sub-components of social cognition that we found to be most cognitively demanding in Chapter 2 (i.e. level 2 perspective-taking, reference assignment, social attention), since they rely more on EF abilities and could therefore benefit from training in these underlying skills. In contrast, we expected that those sub-components of social cognition that we found to be relatively automatic and cognitively efficient (i.e. social inferencing, level 1 perspective-taking, empathy for others' pain) would be less susceptible to enhancement by EF training given the weaker reliance on cognitive skills. Moreover, we predicted that training effects would be greatest among adolescents and older adults since Chapter 2 showed that their social cognition abilities were impaired relative to young adults on those tasks that measured cognitively demanding sub-components, meaning that they have greater capacity for improvement.

4.2. Methods

4.2.1. Participants

From the total sample of 293 participants (see Chapter 2), nine participants were excluded due to low MoCA scores, and 46 were removed as they did not return to complete session two after training, and, six were excluded because they completed less than eight training sessions in the 21-day period. Thus, the final total sample for this training study was 232, divided into three age groups: 71 adolescents (aged 10-19 years), 82 younger adults (aged 20-40 years), and 79 older adults (aged 60-80 years).

Participant details, including mean age, gender balance, SES, IQ, MoCA and AQ-10 for each of the three age groups, are presented in Tables 4. 1 and 4. 2.

Table 4. 1: Participant characteristics by age group (mean values, with standard deviations in parenthesis).

	Adolescents	**Younger Adults**	**Older Adults**
N	71	82	79
Age (years)	14.4 (3.0)	27.6 (5.4)	67.9 (5.2)
Gender (F:M)	39:32	53:29	51:28
SES Index	10.3 (3.4)	10.6 (2.8)	11.4 (2.4)
Full Scale IQ	102.9 (10.5)	101.6 (13.7)	109.8 (11.0)
Verbal IQ	101.5 (9.5)	99.6 (10.9)	107.3 (12.9)
Perceptual Reasoning IQ	105.5 (11.8)	103.4 (11.3)	109.7 (12.6)
MoCA	27.8 (1.9)	27.9 (1.6)	27.2 (1.9)
AQ-10	2.8 (1.9)	3.1 (1.8)	2.6 (1.5)

Table 4. 2: Participant characteristics by age group and training group (mean values, with standard deviations in parenthesis).

Adolescents			**Young**			**Old**		
N	M_{age}	*F:M*	*N*	M_{age}	*F:M*	*N*	M_{age}	*F:M*
Inhibitory control								
18	14.3 (3.1)	11:7	22	29 (5.7)	13:9	20	67.6 (5.2)	15:5
Working memory								
17	14.3 (3.1)	9:8	20	27.0 (5.5)	12:8	20	68.4 (5.9)	13:7
Cognitive flexibility								
17	14.3 (2.9)	11:7	23	26.7 (5.4)	16:7	20	66.8 (4.8)	12:8
Control group								

18	14.5 (2.9)	7:10	18	27.4 (5.4)	12:6	19	69.2 (4.9)	11:8

4.2.2. Experimental tasks

We adopted the same six social cognition tasks employed in Chapter 2. Different versions of the same tasks were used to compare participants' performance at the pre- *versus* post-training sessions in the Director's task and in the Hierarchy of social inferences task. This ensured that participant responses would not be biased by their memory of objects and their locations in the former, or social inferences to make for specific videos in the latter. In addition, we adopted different sets of stimuli for the pre- and post-training sessions for the Empathy task (stimuli were always counterbalanced between physical and social pain/no-pain stimuli), and for the Face to Face conversation we used a different set of poster stimuli behind the experimenter (Averted gaze, Neutral, Sharing Gaze) and asked a different set of questions (see APPENDIX C).

To measure cognitive abilities, we adopted the cognitive tasks described in Chapter 3. However, given debates about the Stroop task as a "pure" measure of IC (Enge, et al., 2014), we replaced this IC assessment task with the Go/No-Go task in this training experiment (see details below).

Go/No-Go (Lustig, 2001). This task was used as a measure of inhibitory control (IC). Individual letter stimuli were presented in the centre of the screen, in black ink on a grey background, and were made up of 21 consonants. Participants were instructed to press the left or right arrow key in response to go and no-go stimuli. Each letter was presented for 300ms followed by a blank screen for 700ms. The task consisted of two practice blocks and two experimental blocks. During the practice blocks, participants received feedback on their accuracy (i.e. a beep indicated an incorrect response), and were required to reach 75% accuracy to proceed. In the first practice block, participants were instructed to press the left arow when an X appeared on the screen (go trials) and to press the right arrow to any other letter (no-go trials). Participants then completed the first experimental block with a total of 200 trials (160 go and 40 no-go trials). In the second practice block, participants were instructed to press the right arrow when an X appeared on the screen (no-go trials) and to press the left arrow to any other letter (go trials). Participants then completed the second experimental block with a total of 200 trials (160 go and 40 no-go trials). This task lasted ~ 6 minutes. The dependent variable was accuracy, calculated as % correct responses to no-go trials, averaged over both experimental blocks.

4.2.3. Procedure

Participants completed the social cognition tasks (described in Chapter 2) and the cognitive assessments tasks (described in Chapter 3 and go/no-go above) in a quiet laboratory at the University of Kent. The entire pre-training session lasted approximately 3.5 hours, and the order of tasks was counterbalanced across participants. Participants were then assigned to a training task (either *n*-back, Stop-signal flanker, task switching, or lexical decision). At the end of the pre-training session, they received instructions on the procedures to complete the online training at home. Participants were invited to complete 10 online training sessions at home, each lasting ~15 minutes, over the next 21 days. We allowed a minimum number of eight and a maximum of 11 training sessions to be completed over this 21-day period (see Table 4. 3 for details on number of sessions performed by each age group and training group).

Training tasks were controlled through INQUISIT software (www.millisecond.com), and participants were sent personalised emails with a link to the appropriate task every two or three days. Following the 21-day training period[5], participants returned to the lab to complete the post-training session, in which they repeated the same six social and three cognitive assessment tasks from the pre-training session, as well as their assigned training tasks.

Table 4. 3: Number of sessions completed by each age group and training group.

Training group	**Adolescents**		**Young**		**Old**	
	Mean	Range	Mean	Range	Mean	Range
Inhibitory control	9.5	8-11	9.6	8-11	9.7	8-11
Working memory	9.2	8-10	9.6	8-11	9.7	8-11
Cognitive flexibility	9.7	8-11	9.7	8-11	9.9	8-11
Control group	9.3	8-11	9.4	8-10	9.7	8-11

4.3. Results

Given the clear improvements in direct transfers found in the previous Chapter, I will focus on near and far transfers from EF training to the EF assessment tasks and the social cognition tasks. In particular, to avoid repetition of age or other effects reported in Chapter 2, discussion of results here will focus on effects that involve training group and time variables. All statistical analyses were conducted in R version 3.6.1.

---

[5] We allowed a minimum of 18 days and a maximum of 25 days between pre- and post-training sessions (*M* = 21.3, *SD* = .90).

4.3.1. Cognitive tasks

A series of ANOVAs were conducted on each of the three cognitive assessment tasks to examine indirect training effects (i.e. near and far transfer), crossing the within-subjects variable Time (pre- *vs*. post-training) with the between-subjects variables Training group (WM *vs.* IC *vs.* CF *vs.* LD) and age Group (adolescents *vs.* young adults *vs.* older adults).

4.3.1.1. OSpan / Working Memory

The final sample in this task was 231 (71 adolescents, 81 young adults, 79 older adults); one participant was excluded due to a technical fault. Partial scores were used as the dependent variable.

As in Chapter 3, results showed a significant effect of Time, $F(3,219) = 13.74$, $p < .001$, $\eta_p^2 = .06$, indicating that participants were more accurate (i.e. had higher partial scores) in the post-training session ($M = 66.0$) compared to the pre-training session ($M = 63.3$). A main effect of Age group was also found $F(2,219) = 8.41$, $p < .001$, $\eta_p^2 = .07$, showing that young adults performed better ($M = 68.0$) than older adults ($M = 60.3$, $t\ (318) = 5.55$, $p < .001$), but did not differ from adolescents ($M = 65.2$, $t\ (302) = 1.80$, $p = .07$). The main effect of Training group was not significant, $F(3, 219) = 1.02$, $p = .39$, $\eta_p^2 = .01$, thus overall recall accuracy was comparable between the four training groups. None of the interactions involving Time, Age group or Training group were significant, Fs < 2.0, ps > .16, suggesting that performance on the OSpan improved from pre- to post-training regardless of training or age group (i.e. most likely reflecting practice).

4.3.1.2. Go/No-Go / Inhibitory Control

For this task, the final sample was 224 (66 adolescents, 80 young adults, 78 older adults); eight participants were excluded due to a technical fault. Percentage correct on no-go trials was used as the dependent variable.

Results revealed a significant main effect of Time, $F(1, 212) = 21.38$, $p < .001$, $\eta_p^2 = .09$, indicating that participants were more accurate at post ($M = .72$) compared to pre-training ($M = .68$). A main effect of Age group was also found, $F(2, 212) = 66.03$, $p < .001$, $\eta_p^2 = .38$, revealing that younger adults performed better ($M = .82$) than older adults ($M = .70$, $t\ (314) = 7.93$, $p < .001$), and adolescents ($M = .57$, $t\ (290) = 7.01$, $p < .001$). The main effect of Training group was not significant, $F(3, 212) = 2.24$, $p = .08$, $\eta_p^2 = .03$, thus overall recall accuracy was comparable between the four training groups. None of the interactions involving Time, Age group or Training group were significant, Fs < 2.5, ps > .09, thus performance on the Go/No-Go task improved from pre- to post-training regardless of training or age group.

4.3.1.3. WCST / Cognitive Flexibility

For this task, the final sample was 219 (64 adolescents, 78 young adults, 77 older adults); 11 participants were excluded due to a technical fault and two participants were excluded due to no correct answers given. Perseverative errors were used as the dependent variable.

As in Chapter 3, results revealed a significant main effect of Time, $F(1, 207) = 14.14$, $p < .001$, $\eta_p^2 = .06$, indicating that participants made significantly fewer preservative errors post-training ($M = 4.91$) compared to pre-training ($M = 6.12$). The main effect of Training group was significant, $F(3, 207) = 2.65$, $p = .05$, $\eta_p^2 = .04$, showing that the IC group ($M = 4.55$) performed significantly better than the LD group (*M* = 6.82, *t* (106) = 2.28, *p* = .02). All other contrasts were non-significant (*t* < 1.57; *p* > .12). A main effect of Age group was also found, *F*(2, 207) = 12.69, *p* < .001, η_p^2 = .11, revealing that young adults performed better (*M* = 3.47) than both older adults (*M* = 6.49, *t*(308) = 4.88, *p* < .001) and adolescents (*M* = 6.84, *t*(308) = 7.16, *p* < .001). Moreover, the Time x Age group interaction was significant, *F* (2, 207) = 3.51, *p* = .03, η_p^2 = .03. Follow-up analyses showed that adolescents (*t*(63) = 2.58, *p* = .01) and older adults (*t*(76) = 3.05, *p* = .003) committed more perseverative errors at pre- compared to post training, but the young adult group showed no difference from pre- to post-training, *t*(77) = .49, *p* = .62. The interactions including Time and Training group were not significant, *F*s < 1.0, *p*s > .40, thus performance on the WCST improved from pre- to post-training regardless of training group.

4.3.2. Hierarchy of Social Inferences

As in Chapter 2, likelihood was calculated as the percentage of trials on which participants responded “yes” that they detected a social behaviour, and reaction times were calculated for trials with a “yes” response. Outliers were excluded from the analysis of reaction times if they fell more than 2.5 standard deviations from the age group’s mean response time and if they were lower than 400ms. The final sample in this task was 225 (67 adolescents, 81 young adults, 77 older adults); seven participants were excluded from the analysis due to a program error. A 3 x 4 x 4 x 2 mixed design ANOVA was used to analyse the likelihood of inferences and log-transformed reaction time data, crossing the between-subjects factors Age Group (adolescents *vs.* young adults *vs.* older adults) and Training Group (WM *vs.* IC *vs.* CF *vs.* LD) with within-subjects variables Inference (Intention, Desire, Belief, Personality) and Time (pre *vs.* post training). Full statistical effects are reported for likelihood in Table 4. 4, and for reaction times in Table 4. 5.

Table 4. 4: Statistical effects for Likelihood in the hierarchy of social inferences task. Asterisks show significance of effects, where ** $p < .01$; *** $p < .001$.

	df	*F*	*p*	η_p^2
Age Group	(2, 213)	16.80	< .001***	.14
Training Group	(3, 213)	2.06	.11	.03
Inference	(3, 213)	200.66	< .001***	.49
Time	(1, 213)	36.20	< .001***	.15
Age Group x Training Group	(6, 213)	3.30	.004**	.09
Age Group x Inference	(6, 639)	1.71	.12	.02
Training Group x Inference	(9, 639)	.53	.85	< .01
Age Group x Time	(2, 213)	1.39	.25	.01
Training Group x Time	(3, 213)	.67	.57	. < .01
Inference x Time	(3, 639)	10.63	< .001***	.05
Age Group x Training Group x Inference	(18, 639)	1.22	.24	.03
Age Group x Training Group x Time	(6, 213)	.61	.72	.02
Age Group x Inference x Time	(6, 639)	1.21	.30	.01
Training Group x Inference x Time	(9, 639)	.35	.96	< .01
Age Group x Training Group x Inference x Time	(18, 639)	.87	.62	.02

Analysis of likelihood revealed a significant main effect of Time, indicating that participants were more likely to detect social behaviour at the pre (M = 78%) compared to post-training session (M = 73%). As in Chapter 2, the main effect of Age Group showed that adolescents (M = 67%) were less likely to make social inferences compared to young adults (M = 77%; $t(146) = 3.61$, $p < .001$), but older adults were more likely to make social inferences (M = 84%) compared to young adults ($t(156) = 2.34$, $p = .02$). In addition, the main effect of Inference showed that participants were more likely to infer a desire (M = 87%) compared to an intention

(*M* = 84%; *t*(224) = 3.43, *p* < .001), more likely to infer an intention compared to a belief (*M* = 74%; *t*(224) = 8.40, *p* < .001), more likely to infer a desire compared to a belief (*t*(224) = 12.01, *p* < .001), and more likely to infer a belief compared to personality (*M* = 57%; *t*(224) = 11.04, *p* < .001).

The significant Inference x Time interaction indicated that overall participants were more likely to infer beliefs at the pre (*M* = 76%) compared to the post training session (*M* = 72%; *t*(224) = 3.36, *p* < .001), and more likely to infer personality at the pre (*M* = 62.6%) compared to the post training session (*M* = 52%; *t*(224) = 6.71, *p* < .001). None of the remaining contrasts reached significance (*ts* < 1.87; *ps* > .06).

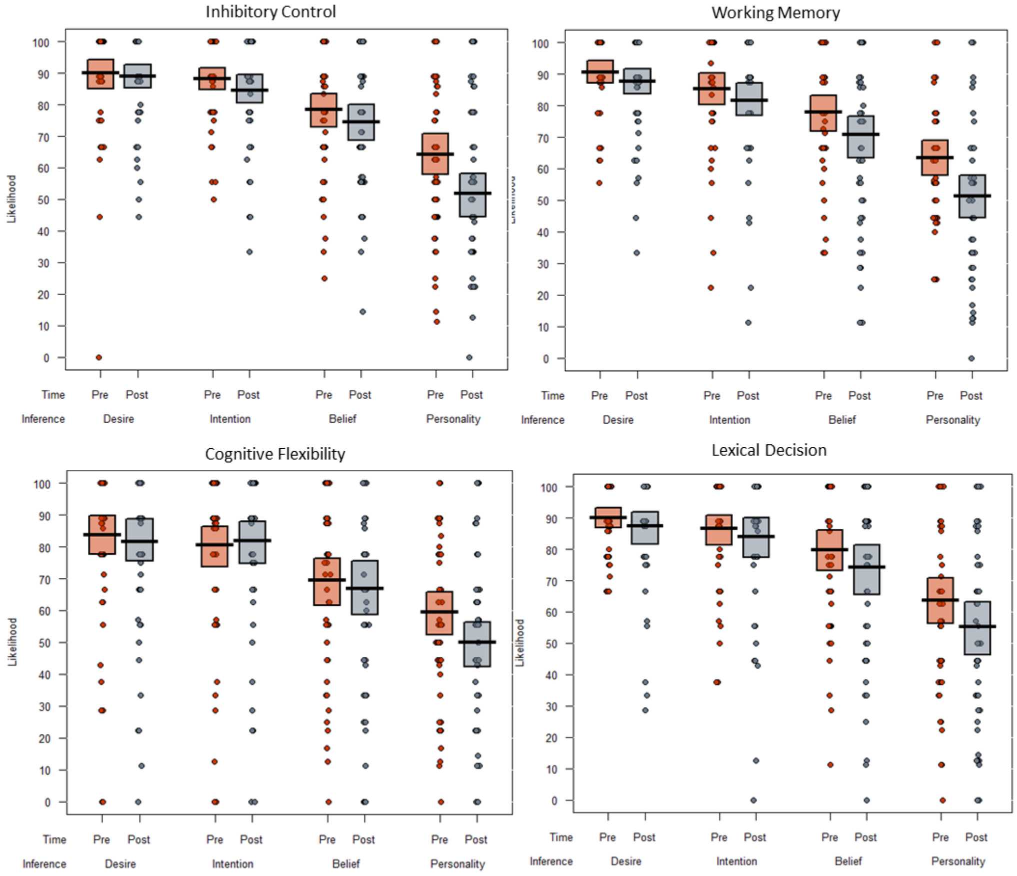


Figure 4. 1: The likelihood of making social inferences for each training group, inference and time (averaged over age group for illustration) in the hierarchy of social inferences task. The plots show raw data points, a horizontal line reflecting the condition mean, and a rectangle representing the Bayesian highest density interval.

Table 4.5: Statistical effects for reaction times in the hierarchy of social inferences task. Asterisks show significance of effects, where ** $p < .01$; *** $p < .001$.

	df	*F*	*p*	η_p^2
Age Group	(2, 213)	28.92	< .001***	.21
Training Group	(3, 213)	.45	.72	< .01
Inference	(3, 213)	40.44	< .001***	.16
Time	(1, 213)	32.27	< .001***	.13
Age Group x Training Group	(6, 213)	1.20	.31	.03
Age Group x Inference	(6, 639)	4.06	< .001***	.04
Training Group x Inference	(9, 639)	.79	.63	.01
Age Group x Time	(2, 213)	2.92	.06	.03
Training Group x Time	(3, 213)	1.12	.34	.02
Inference x Time	(3, 639)	4.46	.004**	.02
Age Group x Training Group x Inference	(18, 639)	.84	.66	.02
Age Group x Training Group x Time	(6, 213)	1.57	.16	.04
Age Group x Inference x Time	(6, 639)	1.39	.22	.01
Training Group x Inference x Time	(9, 639)	1.33	.22	.02
Age Group x Training Group x Inference x Time	(18, 639)	.63	.88	.02

Results on reaction times revealed a main effect of Time, indicating that participants took longer to infer social behaviour at the pre (*M* = 1904ms) compared to post-training session (*M* = 1695ms). As in Chapter 2, the main effect of Age Group showed that older adults were slower to make social inferences (*M* = 2091ms) compared to young adults (*M* = 1705ms; $t(156) = 7.50$, $p < .001$), but young adults and adolescents (*M* = 1743ms) did not differ, $t(146) = .33$, $p = .74$). The main effect of Inference showed that participants were equally fast at inferring a desire (*M* = 1779ms) and an intention (*M* = 1780ms; $t(224) = .48$, $p = .63$), but slower to infer a belief (*M* = 1890ms) compared to either an intention ($t(224) = 6.82$, $p < .001$) or a desire (*t*(224) = 7.30, *p* < .001), and faster to infer a belief compared to a personality (*M* = 1944ms; *t*(224) = 2.71, *p* = .007).

The Age Group x Inference interaction found in Chapter 2 was also replicated here. Importantly, the Inference x Time interaction was significant, indicating that all four social inferences were slower at the pre- compared to post-training session (all *t*s > 2.90, *p*s < .004), though this effect was greater in desire inferences, followed by intentions and belief and smaller when making personality inferences.

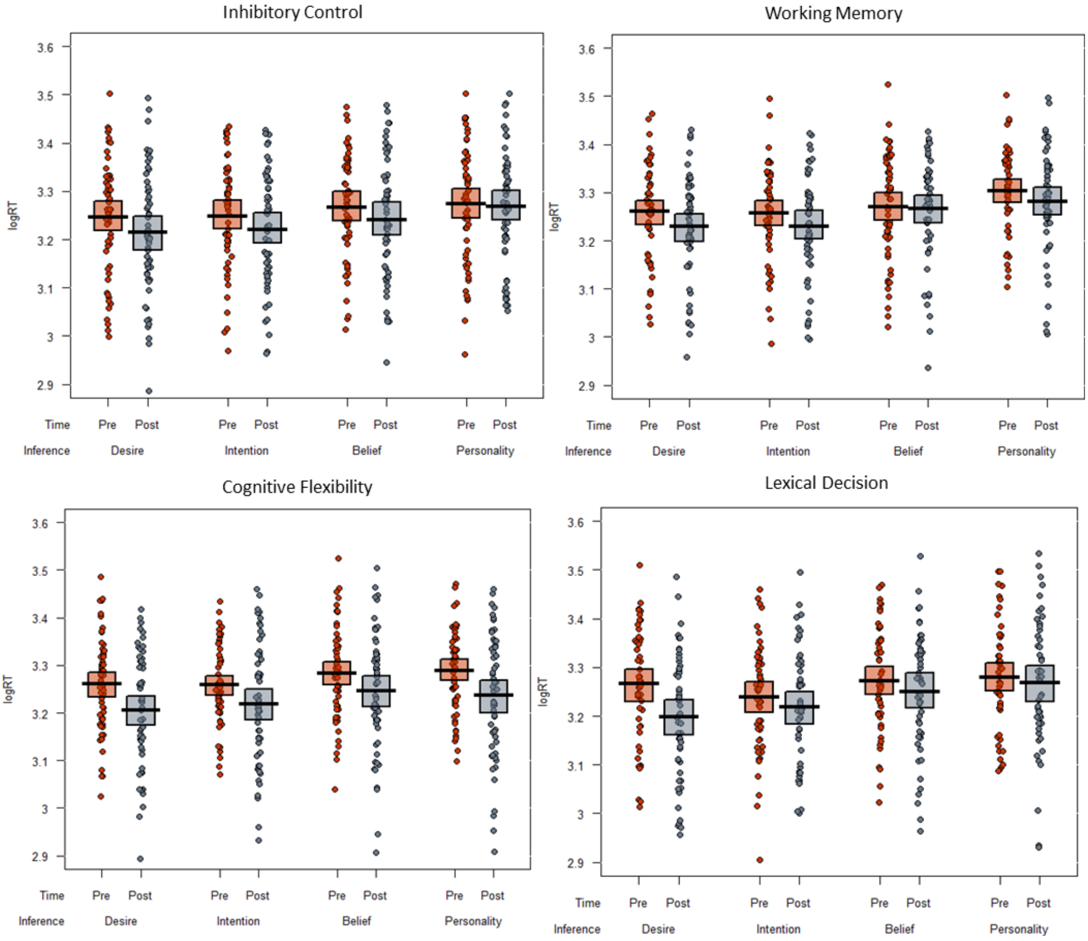


Figure 4. 2: The log reaction time in each training group at the pre and post-training session in the Hierarchy of social inferences task. The plots show raw data points, a horizontal line reflecting the condition mean, and a rectangle representing the Bayesian highest density interval.

4.3.3. Visual/Spatial perspective-taking

As in Chapter 2, only matching trials were included in our analyses, and reaction times were calculated only on correct trials. Outliers were excluded from the analysis of reaction times if they fell more than 2.5 standard deviations from the age group's mean response time. The final sample in this task was 231 (71 adolescents, 81 young adults, 78 older adults); three participants were excluded from the analysis due to a program error. A 3 x 4 x 4 x 2 x 2 x 2 mixed design ANOVA was used for analysis of accuracy and log-transformed reaction time data, crossing the between-subjects factors Age Group (adolescents *vs.* young adults *vs.* older adults) and Training Group (WM *vs.* IC *vs.* CF *vs.* LD), with within-subjects variables Angle (0° *vs.* 60° *vs.* 120° *vs.* 180°), Level (1 *vs.* 2), Content (Visual *vs.* Spatial) and Time (pre *vs.* post training). Full statistical effects for reaction times are reported in Table 4. 6.

Table 4. 6: Statistical effects for reaction times in the visuo-spatial perspective-taking task. Asterisks show significance of effects, where * $p < .05$; ** $p < .01$; *** $p < .001$.

Effect	*df*	*F*	*p*	η_p^2
Age Group	(2, 218)	61.773	<.001***	.36
Training Group	(3, 218)	4.805	.003**	.06
Content	(1, 218)	240.395	<.001***	.52
Angle	(3, 654)	511.081	<.001***	.70
Level	(1, 218)	851.986	<.001***	.80
Time	(1, 218)	283.068	<.001***	.57
Age Group x Training Group	(6, 218)	0.277	0.95	<.01
Age Group x Content	(2, 218)	0.605	0.55	<.01
Training Group x Content	(3, 218)	1.042	0.38	<.01
Age Group x Angle	(6, 654)	11.114	<.001***	.09
Training Group x Angle	(9, 654)	0.432	.92	<.01
Age Group x Level	(2, 218)	3.426	.034*	.03
Training Group x Level	(3, 218)	0.538	0.66	<.01
Age Group x Time	(2, 218)	4.735	0.01**	.04
Training Group x Time	(3, 218)	13.701	<.001***	.16
Content x Angle	(3, 654)	12.501	<.001***	.05
Content x Level	(1, 218)	224.216	<.001***	.51
Angle x Level	(3, 654)	362.554	<.001***	.62

Content x Time	(1, 218)	1.704	.19	<.01
Angle x Time	(3, 654)	3.282	.02*	<.01
Level x Time	(1, 218)	1.038	.31	<.01
Age Group x Training Group x Content	(6, 218)	1.567	.16	<.01
Age Group x Training Group x Angle	(18, 654)	1.82	.02*	.05
Age Group x Training Group x Level	(6, 218)	1.03	.41	.03
Age Group x Training Group x Time	(6, 218)	2.06	.06	.05
Age Group x Content x Angle	(6, 654)	2.45	.02*	.02
Training Group x Content x Angle	(9, 654)	1.20	.29	.02
Age Group x Content x Level	(2, 218)	.50	.61	<.01
Training Group x Content x Level	(3, 218)	.59	.62	<.01
Age Group x Angle x Level	(6, 654)	2.55	.02*	<.01
Training Group x Angle x Level	(9, 654)	1.40	.19	.02
Age Group x Content x Time	(2, 218)	2.04	.13	.02
Training Group x Content x Time	(3, 218)	1.24	.30	.02
Age Group x Angle x Time	(6, 654)	1.37	.23	.01
Training Group x Angle x Time	(9, 654)	.77	065	.01
Age Group x Level x Time	(2, 218)	3.67	.03*	.03
Training Group x Level x Time	(3, 218)	.38	.77	<.01
Content x Angle x Level	(3, 654)	11.25	<.001***	.05
Content x Angle x Time	(3, 654)	3.25	.02*	.02
Content x Level x Time	(1, 218)	5.70	.02*	.03
Angle x Level x Time	(3, 654)	3.85	.009	.02
Age Group x Training Group x Content x Angle	(18, 654)	.92	.55	.03
Age Group x Training Group x Content x Level	(6, 218)	1.79	.10	.05
Age Group x Training Group x Angle x Level	(18, 654)	.38	.99	.01
Age Group x Training Group x Content x Time	(6, 218)	.93	.48	.03
Age Group x Training Group x Angle x Time	(18, 654)	1.35	.15	.04
Age Group x Training Group x Level x Time	(6, 218)	2.68	.02*	.07
Age Group x Content x Angle x Level	(6, 654)	.82	.55	<.01
Training Group x Content x Angle x Level	(9, 654)	1.23	.27	.02
Age Group x Content x Angle x Time	(6, 654)	.57	.75	<.01
Training Group x Content x Angle x Time	(9, 654)	.77	.64	.01
Age Group x Content x Level x Time	(2, 218)	5.04	.007**	.04
Training Group x Content x Level x Time	(3, 218)	.35	.79	<.01
Age Group x Angle x Level x Time	(6, 654)	.61	.72	<.01
Training Group x Angle x Level x Time	(9, 654)	.96	.47	.01
Content x Angle x Level x Time	(3, 654)	4.38	.005**	.02

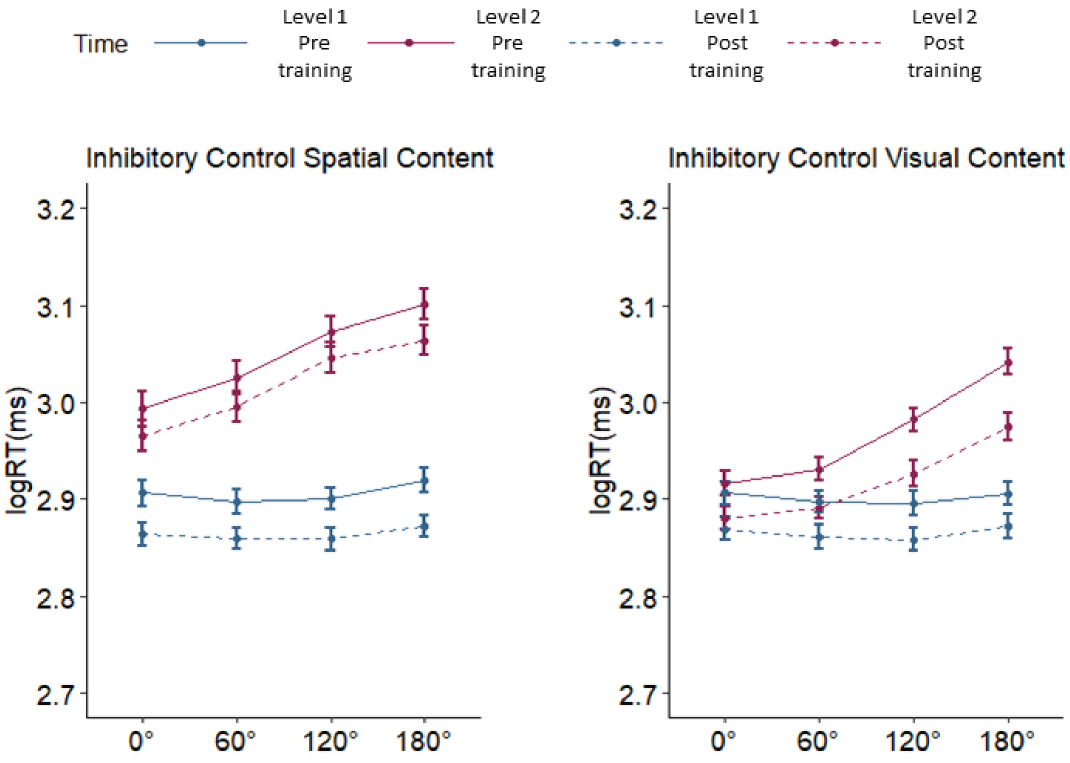
Age Group x Training Group x Content x Angle x Level	(18, 654)	.82	.68	.02
Age Group x Training Group x Content x Angle x Time	(18, 654)	1.08	.37	.03
Age Group x Training Group x Content x Level x Time	(6, 218)	.58	.74	.02
Age Group x Training Group x Angle x Level x Time	(18, 654)	.63	.88	.02
Age Group x Content x Angle x Level x Time	(6, 654)	.75	.61	<.01
Training Group x Content x Angle x Level x Time	(9, 654)	1.31	.23	.02
Age Group x Training Group x Content x Angle x Level x Time	(18, 654)	.604	.90	.02

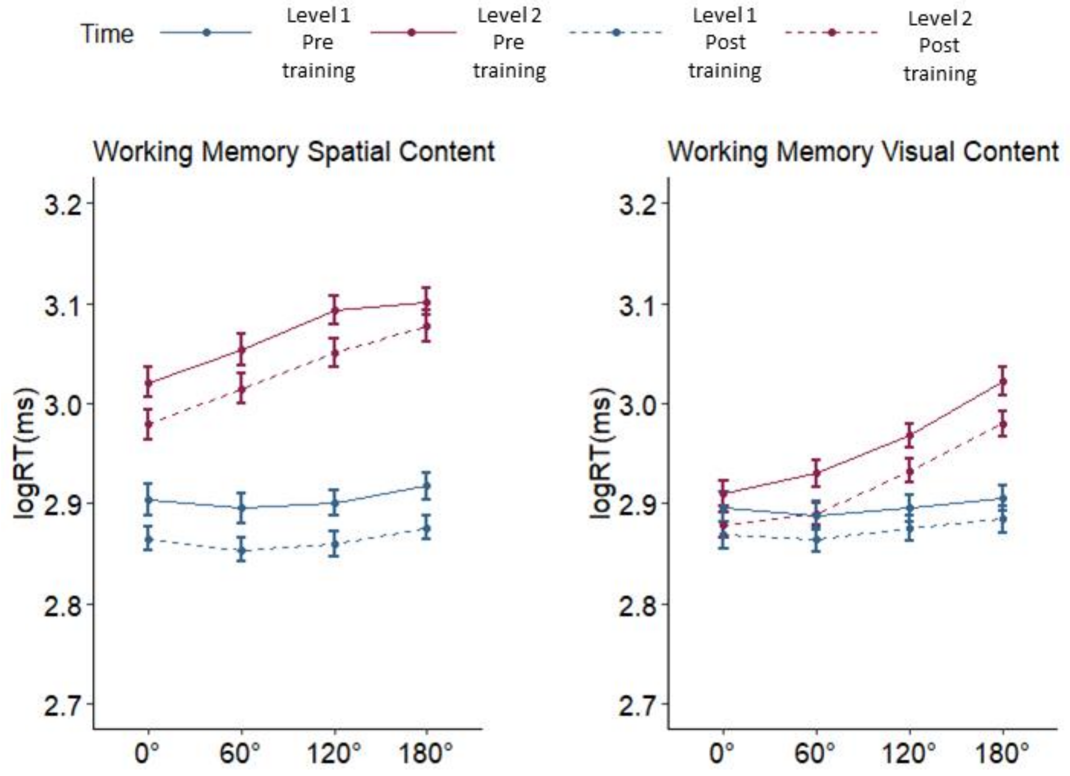
Results revealed significant main effects for all variables. As found in Chapter 2, participants responded slower when making spatial compared to visual judgements (955ms *vs.* 831ms), and when adopting a level 2 perspective compared to a level 1 perspective (1009ms *vs.* 777ms). The main effect of Angle showed that response times increased as the angle of rotation increased (M_{0° = 842; M_{60° = 859; M_{120° = 907; M_{180° = 965; all contrasts $p < .001$). The main effect of Age group showed that older adults spent longer judging the avatar's perspective (1008ms) compared to both adolescents (901ms) and young adults (775ms), and adolescents were slower compared to young adults (all contrasts $p < .001$). The main effect of Training group showed that the cognitive flexibility group judged the avatar's perspective faster (838ms) compared to the control group (M = 932ms, $t(111) = 2.77$, $p = .007$), but none of the other contrasts were significant ($p > .60$). Finally, the main effect of Time indicated that participants were faster in judging the avatar's perspective at the post-training session (842ms) compared to the pre-training session (M = 943ms).

The basic interactions reported in Chapter 2 were replicated here, namely Age Group x Angle, Content x Angle, Content x Level, Angle x Level, Age Group x

Angle x Level, and Content x Angle x Level. Importantly, the interaction between Training group and Time was significant and was further investigated by running one-way ANOVAs to compare training groups separately for pre- and post-training sessions. Results showed that the four training groups differed in performance at the post-training session, $F(3, 226) = 7.06$, $p < .001$, $\eta_p^2 = .09$, but not at the pre-training session, $F(3, 226) = .82$, $p = .48$. Follow-up contrasts between the four groups in the post-training session showed that participants the cognitive flexibility group responded significantly faster (761ms) compared to the inhibitory control group (859ms, $t(118) = 3.54$, $p < .001$), working memory group (869ms, $t(115) = 3.93$, $p < .001$) and lexical decision group (888ms, $t(111) = 3.83$, $p < .001$). None of the other contrasts reached significance (all *ts* < .66, *ps* > .51).

A three-way interaction was also found between Age group, Angle and Training group, however none of the underlying interactions reached significance (all *p*s > .83).


Time
Level 1 Pre training
Level 2 Pre training
Level 1 Post training
Level 2 Post training
Inhibitory Control Spatial Content
Inhibitory Control Visual Content
logRT(ms)
3.2
3.1
3.0
2.9
2.8
2.7
0°
60°
120°
180°


Time
Level 1 Pre training
Level 2 Pre training
Level 1 Post training
Level 2 Post training
Working Memory Spatial Content
Working Memory Visual Content
logRT(ms)
3.2
3.1
3.0
2.9
2.8
2.7
0°
60°
120°
180°

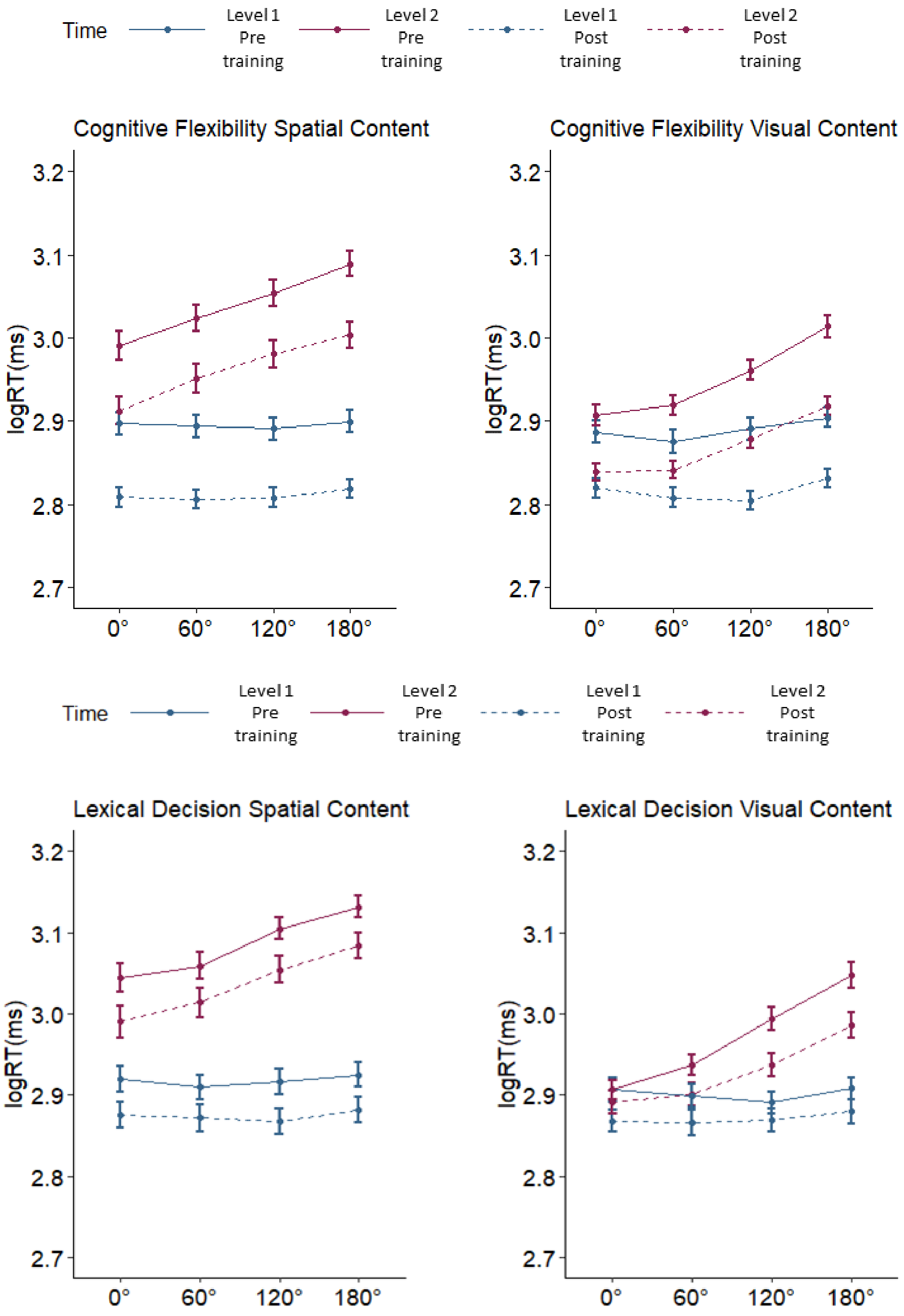


Figure 4. 3: The log reaction time in each condition and training group in the visuo-spatial perspective-taking task at pre and post-training sessions, showing the condition mean and standard errors. Note that the data is collapsed over age groups since this variable did not interact meaningfully with training group, and to facilitate presentation.

4.3.4. Interactive reference assignment

As in Chapter 2, only experimental trials were included in our analyses. Accuracy was calculated as the proportion of trials on which participants correctly selected the target object. Reaction times were calculated for correct responses only, and excluded trials that fell more than 2.5 standard deviations from the age group’s mean response time for each session, or were faster than 200 ms (since this indicated they selected the object before hearing the scalar contrast term). The final sample in this task was 224 participants (67 adolescents, 80 young adults, 77 older adults); ten participants were excluded from the analysis due to a program error. A 3 x 4 x 2 x 2 mixed design ANOVA was used for analysis of accuracy and log-transformed reaction time data, crossing the between-subjects factors Age Group (adolescents *vs.* young adults *vs.* older adults) and Training Group (WM *vs.* IC *vs.* CF *vs*. LD) with the within-subjects factors Condition (listener only *vs.* shared view) and Time (pre *vs.* post training). Full statistical effects are reported in Table 4. 7 for accuracy, and in Table 4. 8 for reaction times.

Analysis of accuracy revealed significant main effects for all variables. The main effect of Time showed that overall participants were more accurate at the post training (*M* = .98) compared to pre training session (*M* = .96). The main effect of Training Group was further investigated by comparing each experimental group (Working Memory, Inhibitory Control and Cognitive Flexibility) with the control group (Lexical Decision), and revealed higher overall accuracy in the cognitive flexibility group (*M* = .98) compared to the control group (*M* = .96, *t*(107) = 2.36, *p* = .02). All other contrasts were non-significant (*p*s > .12). In addition, the main effects of Age Group and Condition replicated effects in Chapter 2 by showing that overall, young adults (*M* = .98) were more accurate at selecting the target object compared to both older adults (*M* = .96; *t*(155) = 3.71, *p* < .001) and adolescents (*M* = .97, *t*(145) = 2.48, *p* = .01), and that participants committed more errors in the listener only condition (*M* = .96) compared to the shared view condition (*M* = .99).

Table 4. 7: Statistical effects for Accuracy in the Director task. Asterisks show significance of effects, where ** *p* < .01, *** *p* < .001.

	df	*F*	*p*	η_p^2
Training Group	(3, 212)	2.80	.04*	.04
Age Group	(2, 212)	11.35	< .001***	.08
Time	(1, 212)	13.16	< .001***	.06
Condition	(1, 212)	33.40	< .001***	.14
Training Group x Age Group	(6, 212)	1,26	.28	.03
Training Group x Time	(3, 212)	1.08	.36	.02
Age Group x Time	(2, 212)	8.73	< .001***	.08
Training Group x Condition	(3, 212)	.83	.48	.01
Age Group x Condition	(2, 212)	7.91	< .001***	.05
Time x Condition	(2, 212)	10.3	.002**	.05
Training Group x Age Group x Time	(6, 212)	.85	.53	.02
Training Group x Age Group x Condition	(6, 212)	1.90	.08	.05
Training Group x Time x Condition	(3, 212)	1.72	.16	.02
Age Group x Time x Condition	(2, 212)	6.42	.002**	.06
Training Group x Time x Condition x Age Group	(6, 212)	1.28	.27	.04

Importantly, the 2-way interactions between Age group x Time and Time x Condition were significant, and were further subsumed under a 3-way interaction between Age Group, Condition and Time. Follow-up analyses showed that the Age Group x Time interaction was significant in the listener only condition, *F*(2, 221) = 8.19, *p* < .001, η_p^2 = .07, but not in the shared view condition, *F*(2, 221) = .59, *p* = .55. To follow up the significant Age group x Time interaction in the Listener only view, we ran one-way ANOVAs to test the effect of Age Group separately for the pre and post training sessions. Results revealed that the effect of Age Group (young adults > adolescents and older adults) was significant in the pre-training session, *F*(2,

221) = 9.64, $p < .001$, $\eta_p^2 = .08$, but not in the post-training session, $F(2, 221) = .70$, showing that the age-differences in accuracy on this task were eliminated after practice.

Table 4.8: Statistical effects for reaction times in the Director task. Asterisks show significance of effects, where *** $p < .001$.

	df	*F*	*p*	η_p^2
Training Group	(3, 212)	.38	.77	< .01
Age Group	(2, 212)	64.16	< .001***	.38
Time	(1, 212)	165.35	< .001***	.44
Condition	(1, 212)	.64	.42	< .01
Training Group x Age Group	(6, 212)	.88	.51	.02
Training Group x Time	(3, 212)	.52	.67	< .01
Age Group x Time	(2, 212)	.11	.90	< .01
Training Group x Condition	(3, 212)	.73	.54	.01
Age Group x Condition	(2, 212)	.47	.63	< .01
Time x Condition	(2, 212)	10.55	.001***	.05
Training Group x Age Group x Time	(6, 212)	1.78	.10	.05
Training Group x Age Group x Condition	(6, 212)	1.53	.17	.04
Training Group x Time x Condition	(3, 212)	.14	.94	< .01
Age Group x Time x Condition	(2, 212)	2.92	.06	.03
Training Group x Time x Condition x Age Group	(6, 212)	.81	.56	.02

Analysis of reaction times revealed significant main effects for Time and Age group. The main effect of Time revealed that participants were slower to select the target object in the pre training (M = 3043ms) compared to post training session (M = 2671ms). The main effect of Age group replicated effects in Chapter 2, showing that older adults were slower to respond (M = 3134ms) compared to young adults (M = 2363ms; $t(155) = 9.46$, $p < .001$), but young adults and adolescents did not differ (M = 2362ms, $t(145) = .68$, $p = .41$).

The Time x Condition interaction was also significant, and was further by comparing participants' reaction times at pre vs post-training session separately for the listener only and shared view conditions. Paired t-tests revealed that participants significantly improved their performance in the listener only condition ($t(223) = 3.46$, $p < .001$) but not in the shared view condition ($t(223) = .92$, $p = .36$).

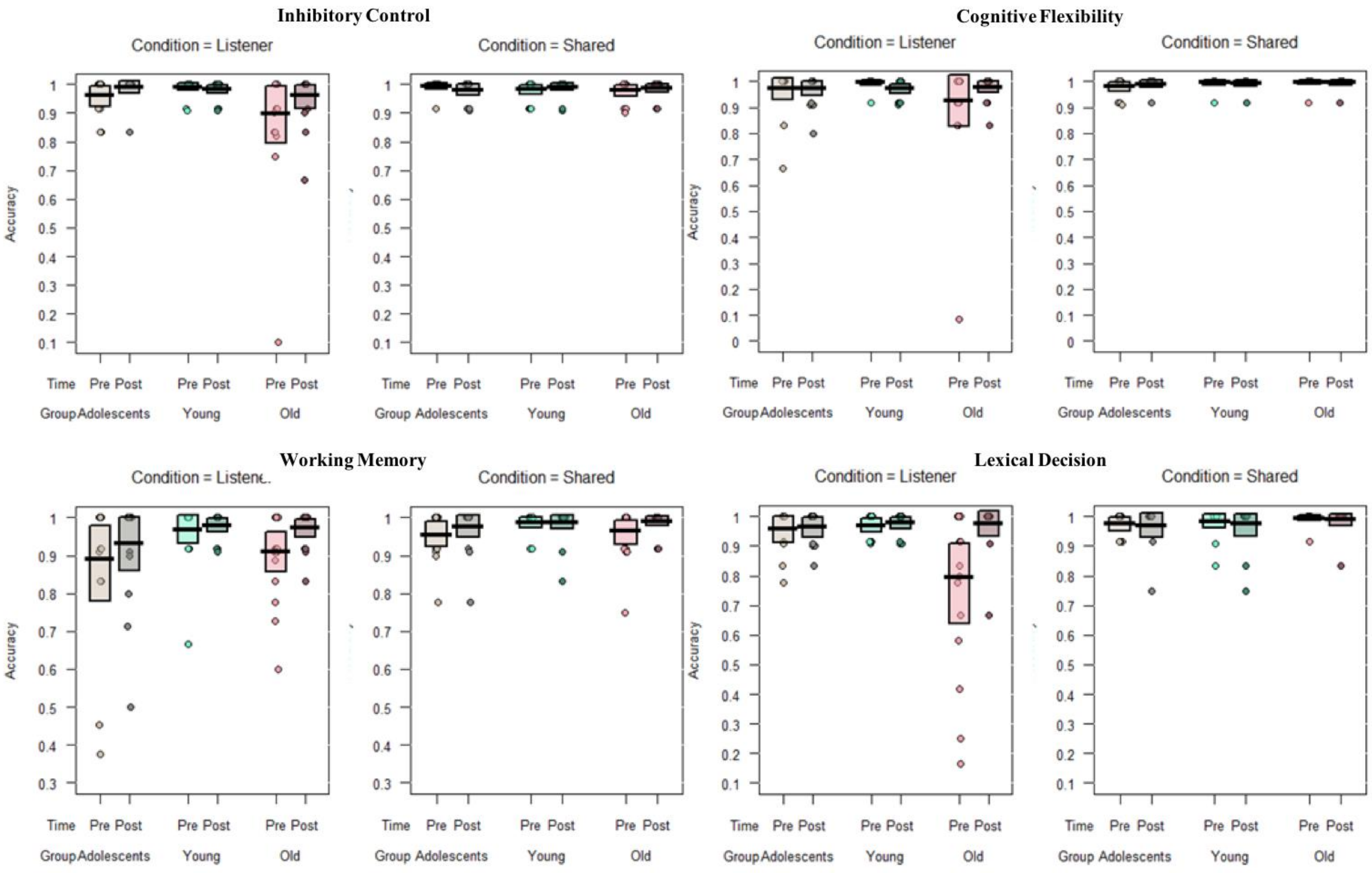


Figure 4.4: The accuracy in each condition per each Age group, Training group at the pre and post-training session in the Director task. The plots show raw data points, a horizontal line reflecting the condition mean, and a rectangle representing the Bayesian highest density interval.

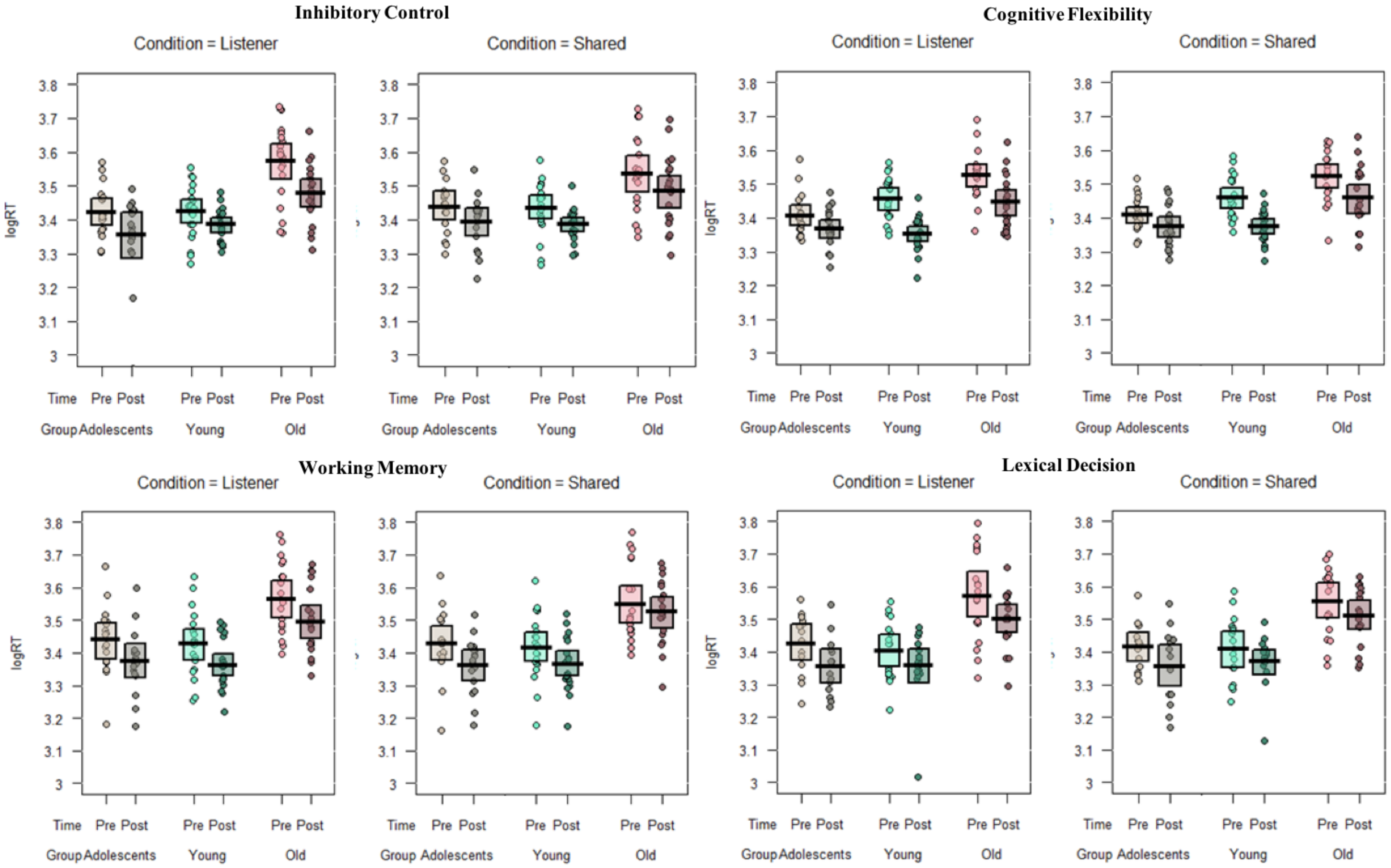


Figure 4. 5: Response times in each condition per each Age group, Training group at the pre and post-training session in the Director task. The plots show raw data points, a horizontal line reflecting the condition mean, and a rectangle representing the Bayesian highest density interval.

4.3.5. Empathy for physical and social pain

The final sample in this task was 175 (46 adolescents, 64 young adults, 65 older adults). From the original sample 13 participants were excluded due to too few segments (more than 25%) and 17 participants were excluded due to computer failure.

4.3.5.1. Pain ratings

Pain ratings were analysed using a 3 x 4 x 2 x 2 x 2 mixed design ANOVA, crossing the between-subjects variable age Group (adolescents *vs.* young adults *vs.* older adults) and Training Group (WM *vs.* IC *vs.* CF *vs.* LD) with the within-subjects variables Type (pain *vs.* no pain), Content (physical *vs.* social) and Time (pre *vs.* post training). Full statistical effects are reported for pain ratings in Table 4. 9.

Table 4. 9: Statistical effects for pain ratings in the empathy for others' pain task. Asterisks show significance of effects, where * $p < .05$; *** $p < .001$.

	df	*F*	*p*	η_p^2
Training Group	(3, 162)	1.04	.38	.02
Age Group	(2, 162)	1.04	.35	.01
Type	(1, 162)	927.22	< .001***	.86
Content	(1, 162)	40.54	< .001***	.20
Time	(1, 162)	10.71	.001***	.06
Training Group x Age Group	(6, 162)	.93	.47	.03
Training Group x Type	(3, 162)	2.71	.05*	.05
Age Group x Type	(2, 162)	2.63	.08	.03
Training Group x Content	(3, 162)	.33	.80	< .01
Age Group x Content	(2, 162)	3.99	.02*	.05
Training Group x Time	(3, 162)	2.43	.07	.04
Age Group x Time	(2, 162)	.51	.60	< .01
Type x Content	(1, 162)	39.33	< .001***	.20
Type x Time	(1, 162)	13.19	< .001***	.08
Content x Time	(1, 162)	1.99	.16	.01
Training Group x Age Group x Type	(6, 162)	1.02	.41	.04
Training Group x Age Group x Content	(6, 162)	1.93	.08	.07

Training Group x Age Group x Time	(6, 162)	.67	.68	.02
Training Group x Type x Content	(3, 162)	1.10	.35	.02
Age Group x Type x Content	(2, 162)	.50	.61	< .01
Training Group x Type x Time	(3, 162)	1.54	.20	.03
Age Group x Type x Time	(2, 162)	1.23	.29	.02
Training Group x Content x Time	(3, 162)	.26	.85	< .01
Age Group x Content x Time	(2, 162)	3.66	.03*	.04
Type x Content x Time	(1, 162)	1.23	.27	< .01
Training Group x Age Group x Type x Content	(6, 162)	1.06	.39	.04
Training Group x Age Group x Type x Time	(6, 162)	.77	.60	.03
Training Group x Age Group x Content x Time	(6, 162)	.51	.80	.02
Training Group x Type x Content x Time	(3, 162)	.69	.56	.01
Age Group x Type x Content x Time	(2, 162)	.33	.71	< .01
Training Group x Age Group x Type x Content x Time	(6, 162)	.21	.97	< .01

As in Chapter 2, the Age Group x Content and Type x Content interactions were significant here. Importantly, the analysis also revealed a significant Time x Type interaction, which showed that participants reduced their pain ratings for no-pain images from the pre (M = 14.6) to post-training session (M = 9.57, $t(173) = 4.76$, $p < .001$), but did not differ from pre to post training when rating pain images ($t(173) = .59$, $p = .56$). Finally, the Age Group x Content x Time interaction was significant. Follow-up analyses tested effects of Content and Time separately for each age group, and revealed that the Content x Time interaction was significant in the adolescent group, $F(1, 43) = 8.49$, $p < .001$, $\eta_p^2 = .17$, but not in the young or older adult groups (Fs < 1.67, ps > .20). To follow up the significant interaction in the adolescents, we used t-tests to compare pain ratings at pre *vs.* post-training separately for the physical and social content. Results showed that adolescents judged physical images as more painful at pre (*M* = 39.2) than post-training (*M* = 35.0, $t(43) = 2.89$, $p < .001$), but did not differ from pre to post training when judging social images, $t(43) = .83$, $p = .41$.

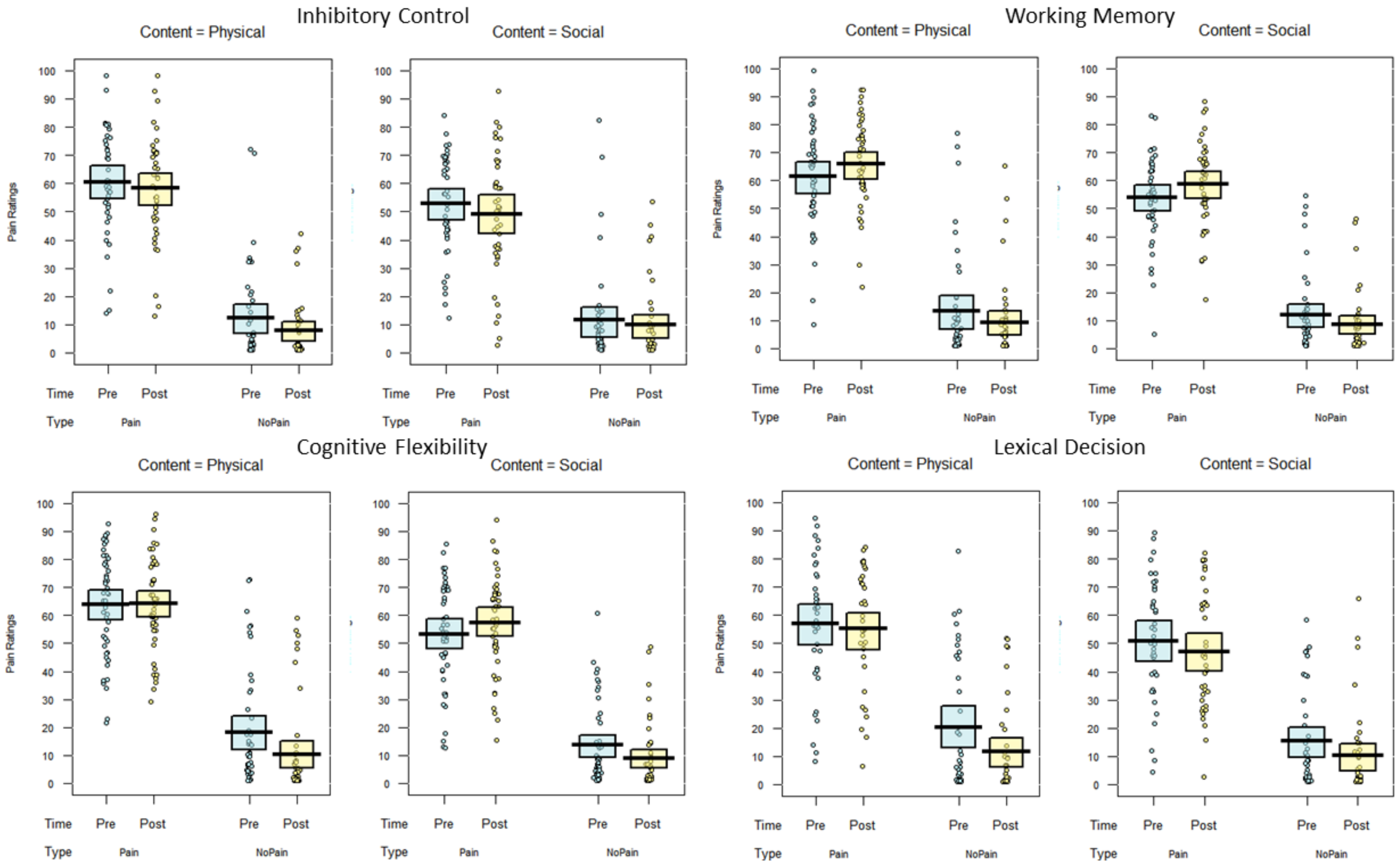


Figure 4. 6: The pain ratings for each training group, condition and time (averaged over age group for illustration) in the empathy for others' pain task. The plots show raw data points, a horizontal line reflecting the content mean, and a rectangle representing the Bayesian highest density interval.

4.3.5.2. EEG suppression

Mu/alpha (8-13Hz) and beta (13-35Hz) suppression was analysed using separate 3 x 4 x 2 x 2 x 2 x 2 mixed ANOVAs that crossed the between-subjects factors Age Group (adolescents *vs.* young adults *vs.* older adults) and Training Group (WM *vs.* IC *vs.* CF *vs.* LD) with the within-subjects factors Type (pain *vs.* no-pain), Content (physical *vs.* social), Electrode site (central *vs.* occipital) and Time (pre *vs.* post training). Full statistical effects are reported in Table 4. 10.

<u>Table 4.10:</u> Statistical effects for mu/alpha and beta suppression in the empathy for others' pain task. Asterisks show significance of effects, where * $p < .05$; ** $p < .01$; *** $p < .001$.

		df	*F*	*p*	η_p^2
Mu/Alpha suppression	Training Group	(3, 163)	1.42	.24	.03
	Age Group	(2, 163)	6.08	.003**	.07
	Time	(1, 163)	21.31	< .001***	.12
	Type	(1, 163)	42.30	< .001***	.20
	Content	(1, 163)	16.40	< .001***	.10
	Electrode	(1, 163)	17.80	< .001***	.10
	Training Group x Age Group	(6, 163)	1.01	0.45	.03
	Training Group x Time	(3,163)	.63	.64	.01
	Age Group x Time	(2, 163)	.71	.69	< .01
	Training Group x Type	(3,163)	.15	.82	< .01
	Age Group x Type	(2, 163)	4.48	.03*	.04
	Training Group x Content	(3,163)	.15	.94	< .01
	Age Group x Content	(2, 163)	.32	.77	< .01
	Training Group x Electrode	(3,163)	1.27	.37	.02
	Age Group x Electrode	(2, 163)	16.60	< .001***	.18
	Time x Type	(1, 163)	3.16	.15	.01
	Time x Content	(1, 163)	.26	.55	< .01
	Type x Content	(1, 163)	.31	.52	< .01
	Time x Electrode	(1, 163)	8.16	.005**	.05
	Type x Electrode	(1, 163)	8.07	.005**	.05
	Content x Electrode	(1, 163)	.40	.53	< .01
	Training Group x Age Group x Time	(6, 163)	.42	.86	.02
	Training Group x Age Group x Type	(6, 163)	.88	.51	.03
	Training Group x Age Group x Content	(6, 163)	1.49	.18	.05
	Training Group x Age Group x Electrode	(6, 163)	.25	.95	< .01
	Training Group x Time x Type	(3, 163)	.25	86	< .01
	Age Group x Time x Type	(2, 163)	.83	.44	.01
	Training Group x Time x Content	(3, 163)	.51	.68	< .01
	Age Group x Time x Content	(2, 163)	.23	.80	< .01

Training Group x Content x Type	(3, 163)	.61	.61	.01
Age Group x Type x Content	(2, 163)	.06	.94	< .01
Training Group x Time x Electrode	(3, 163)	1.32	.27	.02
Age Group x Time x Electrode	(2, 163)	2.74	.07	.03
Training Group x Type x Electrode	(3, 163)	2.58	.06	.05
Age Group x Type x Electrode	(2, 163)	.24	.79	< .01
Training Group x Content x Electrode	(3, 163)	.97	.40	.02
Age Group x Content x Electrode	(2, 163)	1.45	.24	.02
Type x Time x Content	(1, 163)	1.20	.28	< .01
Electrode x Type x Time	(1, 163)	.72	.40	< .01
Time x Content x Electrode	(1, 163)	1.26	.26	< .01
Type x Content x Electrode	(1, 163)	.002	.97	< .01
Training Group x Time x Age Group x Type	(6, 163)	1.35	.24	.05
Training Group x Time x Age Group x Content	(6, 163)	.50	.81	.02
Training Group x Age Group x Type x Content	(6, 163)	.59	.74	.02
Training Group x Age Group x Electrode x Time	(6, 163)	.66	.68	.02
Training Group x Age Group x Type x Electrode	(6, 163)	1.29	.27	.05
Training Group x Age Group x Content x Electrode	(6, 163)	1.13	.35	.04
Training Group x Type x Content x Time	(3, 163)	.15	.93	< .01
Age Group x Time x Type x Content	(2, 163)	.20	.82	< .01
Training Group x Type x Electrode x Time	(3, 163)	.48	.69	< .01
Training Group x Time x Content x Electrode	(3, 163)	.70	.55	.01
Age Group x Time x Content x Electrode	(2, 163)	.44	.64	< .01
Training Group x Type x Content x Electrode	(3, 163)	.40	.75	< .01
Age Group x Type x Content x Electrode	(2, 163)	4.14	.02*	.05
Time x Type x Content x Electrode	(1, 163)	.02	.89	< .01

	Training Group x Age Group x Time x Type x Content	(6, 163)	1.75	.11	.06
	Training Group x Age Group x Time x Type x Electrode	(6, 163)	.49	.81	.02
	Training Group x Age Group x Time x Content x Electrode	(6, 163)	.53	.79	.02
	Training Group x Age Group x Type x Content x Electrode	(6, 163)	.27	.95	.01
	Training Group x Time x Type x Content x Electrode	(3, 163)	.37	.77	< .01
	Age Group x Time x Type x Content x Electrode	(2, 163)	.12	.89	< .01
	Training Group x Age Group x Time x Type x Content x Electrode	(6, 163)	.60	.73	.02
Beta suppression	Training Group	(3, 163)	1.77	.15	.03
	Age Group	(2, 163)	10.33	< .001***	.11
	Time	(1, 163)	9.66	< .001***	.06
	Type	(1, 163)	21.82	< .001***	.12
	Content	(1, 163)	15.25	< .001***	.09
	Electrode	(1, 163)	4.91	.02*	.03
	Training Group x Age Group	(6, 163)	.27	.95	.01
	Training Group x Time	(3,163)	1.04	.38	.02
	Age Group x Time	(2, 163)	1.60	.20	.02
	Training Group x Type	(3,163)	.27	.85	< .01
	Age Group x Type	(2, 163)	5.12	.007**	.06
	Training Group x Content	(3,163)	.68	.57	.01
	Age Group x Content	(2, 163)	4.02	.02*	.05
	Training Group x Electrode	(3,163)	1.30	.28	.02
	Age Group x Electrode	(2, 163)	.70	.50	< .01
	Time x Type	(1, 163)	.68	.41	< .01
	Time x Content	(1, 163)	1.36	.25	< .01
	Type x Content	(1, 163)	.11	.74	< .01
	Time x Electrode	(1, 163)	2.50	.12	.02
	Type x Electrode	(1, 163)	5.17	.02*	.03
	Content x Electrode	(1, 163)	7.36	.007**	.04
	Training Group x Age Group x Time	(6, 163)	.86	.53	.03
	Training Group x Age Group x Type	(6, 163)	.63	.70	.02

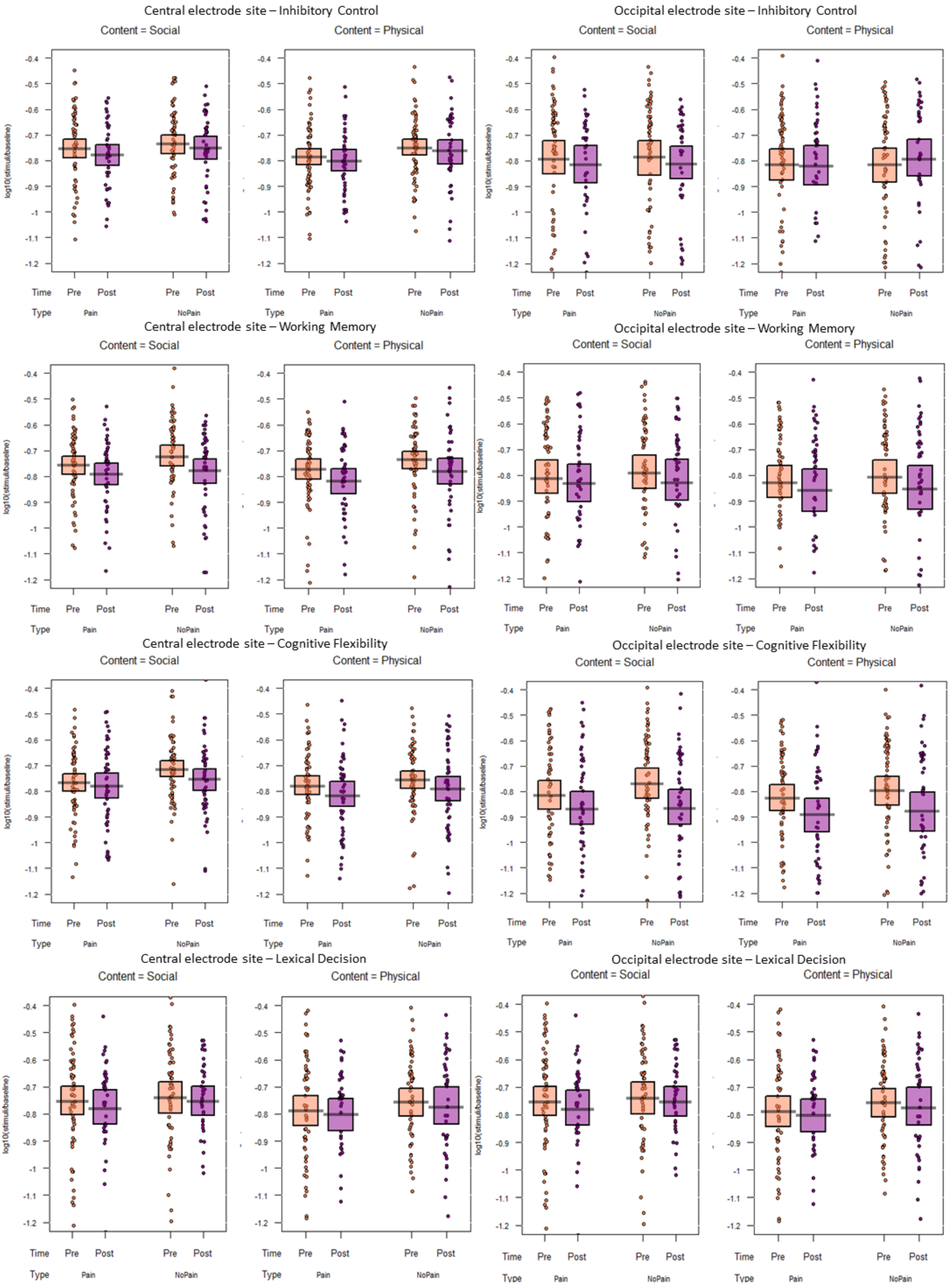
Training Group x Age Group x Content	(6, 163)	.55	.77	.02
Training Group x Age Group x Electrode	(6, 163)	1.0	.43	.04
Training Group x Time x Type	(3, 163)	.12	.95	< .01
Age Group x Time x Type	(2, 163)	.15	.86	< .01
Training Group x Time x Content	(3, 163)	.30	.83	< .01
Age Group x Time x Content	(2, 163)	4.70	.01**	.06
Training Group x Content x Type	(3, 163)	.66	.58	.01
Age Group x Type x Content	(2, 163)	1.22	.30	.02
Training Group x Time x Electrode	(3, 163)	1.52	.21	.03
Age Group x Time x Electrode	(2, 163)	1.21	.30	.02
Training Group x Type x Electrode	(3, 163)	2.08	.11	.04
Age Group x Type x Electrode	(2, 163)	.51	.60	< .01
Training Group x Electrode x Content	(3, 163)	.83	.48	.02
Age Group x Content x Electrode	(2, 163)	2.00	.14	.02
Time x Type x Content	(1, 163)	.65	.42	< .01
Time x Type x Electrode	(1, 163)	.28	.60	< .01
Time x Content x Electrode	(1, 163)	.07	.79	< .01
Type x Content x Electrode	(1, 163)	1.51	.22	< .01
Training Group x Time x Age Group x Type	(6, 163)	1.30	.26	.05
Training Group x Time x Age Group x Content	(6, 163)	.21	.97	< .01
Training Group x Type x Age Group x Content	(6, 163)	1.03	.41	.04
Training Group x Age Group x Time x Electrode	(6, 163)	.71	.64	.03
Training Group x Age Group x Electrode x Type	(6, 163)	.24	.96	< .01
Training Group x Age Group x Content x Electrode	(6, 163)	1.75	.12	.06
Training Group x Time x Content x Type	(3, 163)	.44	.73	< .01
Age Group x Type x Content x Time	(2, 163)	.89	.41	.01
Training Group x Time x Type x Electrode	(3, 163)	1.26	.29	.02

Age Group x Type x Electrode x Time	(2, 163)	.61	.55	< .01
Training Group x Time x Content x Electrode	(3, 163)	.18	.91	< .01
Age Group x Time x Content x Electrode	(2, 163)	.21	.82	< .01
Training Group x Type x Content x Electrode	(3, 163)	2.46	.07	.04
Age Group x Type x Content x Electrode	(2, 163)	.99	.37	.01
Time x Type x Content x Electrode	(1, 163)	.49	.49	< .01
Training Group x Age Group x Time x Type x Content	(6, 163)	.41	.87	.02
Training Group x Age Group x Time x Type x Electrode	(6, 163)	1.75	.11	.06
Training Group x Age Group x Time x Content x Electrode	(6, 163)	1.67	.13	.06
Training Group x Age Group x Type x Content x Electrode	(6, 163)	.29	.94	.01
Training Group x Time x Type x Content x Electrode	(3, 163)	1.07	.36	.02
Age Group x Time x Type x Content x Electrode	(2, 163)	.22	.80	< .01
Training Group x Age Group x Time x Type x Content x Electrode	(6, 163)	.63	.71	.02

4.3.5.2.1. Mu/alpha suppression

Analysis of mu/alpha suppression replicated the main effects reported in Chapter 2. Overall mu/alpha suppression was greater for images that depicted pain (*M* = -80) compared to no-pain (*M* = -77), for physical (*M* = -.80) compared to social stimuli (*M* = -.78), and over the occipital (*M* = -.80) compared to central electrodes (*M* = -.77). The main effect of Age Group was also replicated here, reflecting greater mu/alpha suppression among young adults (*M* = -.82) compared to older adults (*M* = -.74, *t*(127) = 3.30, *p* = .001), but no difference between young adults and adolescents (*M* = -.80, *t*(108) = -.54, *p* = .59). In addition, the main effect of Time was significant, showing that overall mu/alpha suppression was greater at the post (*M* = -.81) compared to the pre-training session (*M* = -.77).

The significant Age Group x Type and Age Group x Electrode interactions reported in Chapter 2 were replicated here. In addition, the 2-way Time x Electrode interaction was significant, revealing increased mu/alpha suppression at the post compared to pre-training session over central (*M* = -.78 *vs.* -.75, *t*(174) = 3.11, *p* = .002) and occipital electrodes (*M* = -.83 *vs.* -.79, *t*(174) = 4.86, *p* < .001), though this difference was greater over occipital than central electrodes.



Central electrode site – Inhibitory Control

Occipital electrode site – Inhibitory Control

Central electrode site – Working Memory

Occipital electrode site – Working Memory

Central electrode site – Cognitive Flexibility

Occipital electrode site – Cognitive Flexibility

Central electrode site – Lexical Decision

Occipital electrode site – Lexical Decision

Figure 4. 7: Mu/alpha suppression for each electrode site and condition in each age group in the empathy for others' pain task. The plots show raw data points, a horizontal line reflecting the content mean, and a rectangle representing the Bayesian highest density interval.

4.3.5.2.2. Beta suppression

Analysis of beta suppression replicated all the main effects reported in Chapter 2. Overall beta suppression was greater for images that depicted pain (*M* = -74) than no-pain (*M* = -73), for physical (*M* = -.74) than social stimuli (*M* = -.73), and over the occipital (*M* = -.74) compared to central electrodes (*M* = -.73). The main effect of Age group was also replicated here, showing greater beta suppression among young adults (*M* = -.76) compared to older adults (*M* = -.70, *t*(127) = 3.97, $p < .001$), but no difference between young adults and adolescents (*M* = -.76, *t*(108) = -.06, $p = .95$). In addition, the significant main effect of Time indicated greater beta suppression at the post (*M* = -.75) compared to pre-training session (*M* = -.73).

The significant Age Group x Type interaction reported in Chapter 2 was replicated here. In addition, the 3-way Time x Age Group x Content interaction was significant. Follow-up analyses tested effects of Content and Time separately for each age group, and revealed that the Content x Time interaction was significant in the young group, $F(1, 63) = 8.68$, $p < .005$, $\eta_p^2 = .12$, but not in the adolescents or older adult groups (*F*s < 2.01, *p*s > .16). To follow up the significant interaction in the young group, we used t-tests to compare beta suppressions at pre *vs.* post-training separately for the physical and social content. Results showed greater beta suppression at pre (*M* = 39.2) than post-training (*M* = 35.0, *t*(63) = 2.32, $p = .02$) in response to social stimuli, but no changes in beta suppression were detected in response to physical stimuli, *t*(63) = .60, $p = .55$.

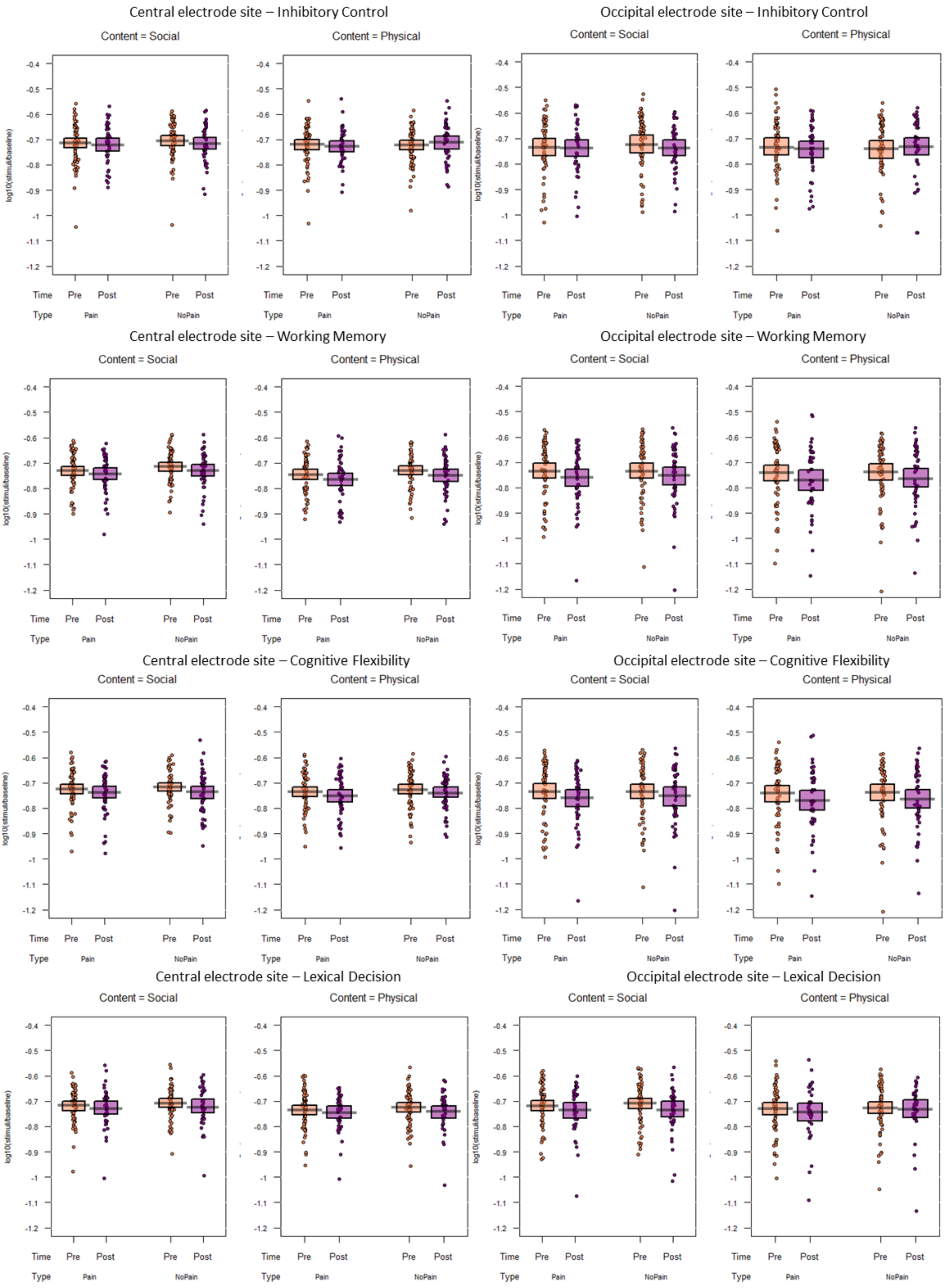


Figure 4. 8: The pain ratings for each training group, condition and time (averaged over age group for illustration) in the empathy for others' pain task. The plots show raw data points, a horizontal line reflecting the content mean, and a rectangle representing the Bayesian highest density interval

4.3.6. Face to face conversation

4.3.6.1. Face to face conversation

The final sample in this task was 191 (60 adolescents, 71 young adults, 60 older adults); 25 participants were excluded due to insufficient eye-tracking data and 14 participants were excluded due to technical issues in one of their testing sessions. Fixations were analysed using a 3 x 4 x 2 x 2 x 3 mixed design ANOVA, crossing the between-subjects factors Age Group (adolescents *vs.* young adults *vs.* older adults) and Training Group (WM *vs.* IC *vs.* CF *vs.* LD) with the within-subjects factors Condition (speaking *vs.* listening), Time (pre *vs.* post training) and AoI (face *vs.* body *vs.* background). Full statistical effects are reported in Table 4. 11[6].

Table 4. 11: Statistical effects for fixations in the face to face conversation task. Asterisks show significance of effects, where *** p < .001.

Effect	*df*	*F*	*p*	η_p^2
Age Group	(2, 179)	<.001	1	<.01
Training Group	(3, 179)	.001	1	<.01
Condition	(1, 179)	.01	1	<.01
AoI	(2, 358)	278.59	< .001***	.61
Time	(1, 179)	.005	1	<.01
Age Group x Training Group	(6, 179)	.002	1	<.01
Age Group x AoI	(4, 358)	6.35	< .001***	.06
Training Group x AoI	(6, 358)	.67	.67	.01
Age Group x Condition	(2, 179)	.01	.99	<.01
Training Group x Condition	(3, 179)	.005	1	<.01
Age Group x Time	(2, 179)	.003	1	<.01
Training Group x Time	(3, 179)	<.001	1	<.01
Condition x AoI	(2, 358)	151.11	< .001***	.46
AoI x Time	(2, 358)	1.12	.33	<.01
Condition x Time	(1, 179)	.10	.75	<.01
Age Group x Training Group x AoI	(12, 358)	1.19	.29	.04

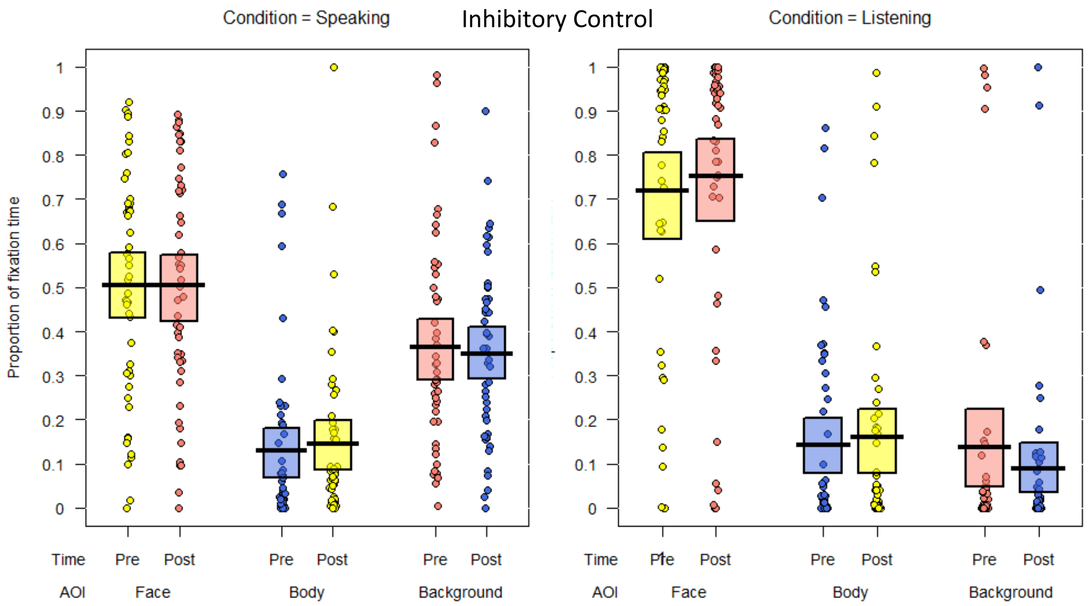
---

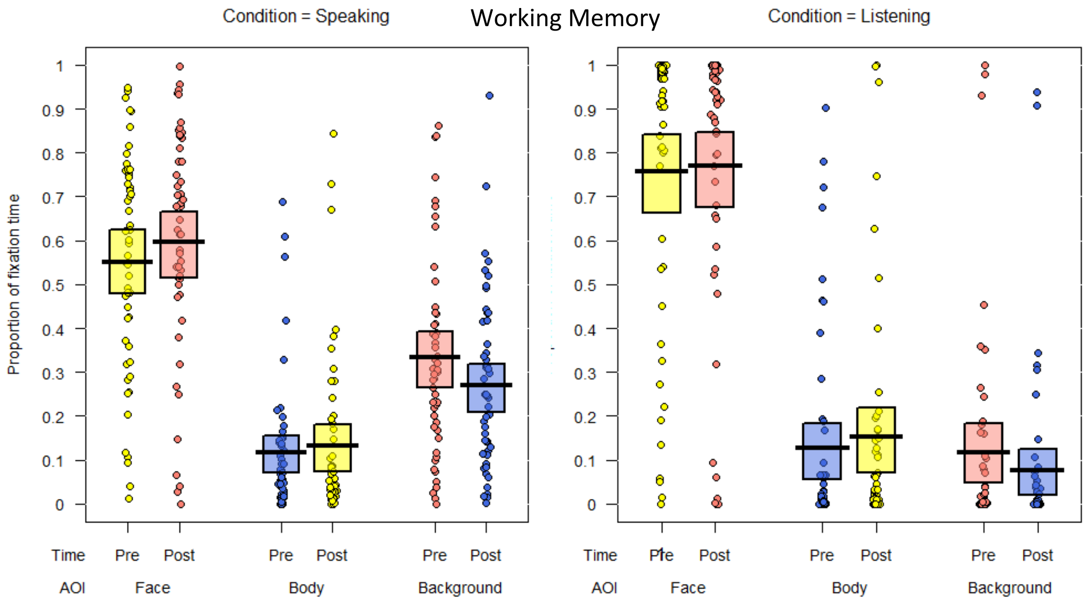
[6] Note that some of the effects are not meaningful in this analysis because proportions of fixations for each participant/condition/time summed to 1.

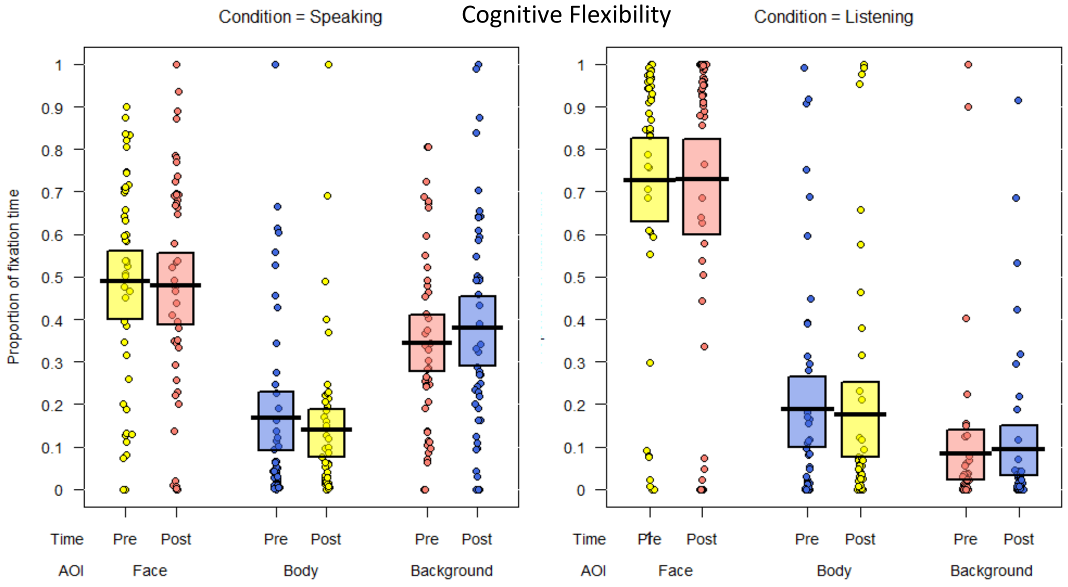
Age Group x Training Group x Time	(6, 179)	.005	1	<.01
Age Group x Training Group x Condition	(6, 179)	.002	1	<.01
Age Group x Condition x AoI	(4, 358)	5.43	< .001***	.06
Training Group x Condition x AoI	(6, 358)	1.38	.22	.02
Age Group x AoI x Time	(4, 358)	.74	.57	<.01
Training Group x AoI x Time	(6, 358)	.82	.55	.01
Age Group x Condition x Time	(2, 179)	.06	.94	<.01
Training Group x Condition x Time	(3, 179)	.07	.98	<.01
AoI x Condition x Time	(2, 358)	.49	.61	<.01
Age Group x Training Group x Condition x AoI	(12, 358)	.27	.99	.009
Age Group x Training Group x AoI x Time	(12, 358)	1.18	.29	.04
Age Group x Training Group x Condition x Time	(6, 179)	.03	1	<.01
Age Group x AoI x Condition x Time	(4, 358)	2.77	.03*	.03
Training Group x AoI x Condition x Time	(6, 358)	.33	.92	<.01
Age Group x Training Group x AoI x Condition x Time	(12, 358)	1.19	.29	.04

Analyses replicated the main effect of AoI, and interactions between Age Group x AoI, Condition x AoI, and Age Group x AoI x Condition reported in Chapter 2, however none of these effects were modulated by Time or Training Group.

The 4-way interaction Age group x AoI x Condition x Time was also significant and further explored, running three separate 3-way ANOVAs on each AoI, however none of the ANOVAs revealed a significant interaction Age group x Time x Condition.


Condition = Speaking
Inhibitory Control
Condition = Listening
Proportion of fixation time
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Time
Pre
Post
Pre
Post
Pre
Post
AOI
Face
Body
Background
Time
Pre
Post
Pre
Post
Pre
Post
AOI
Face
Body
Background


Condition = Speaking
Working Memory
Condition = Listening
Proportion of fixation time
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Time
Pre
Post
Pre
Post
Pre
Post
AOI
Face
Body
Background
Time
Pre
Post
Pre
Post
Pre
Post
AOI
Face
Body
Background


Condition = Speaking
Cognitive Flexibility
Condition = Listening
Proportion of fixation time
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Time
Pre
Post
Pre
Post
Pre
Post
AOI
Face
Body
Background
Time
Pre
Post
Pre
Post
Pre
Post
AOI
Face
Body
Background

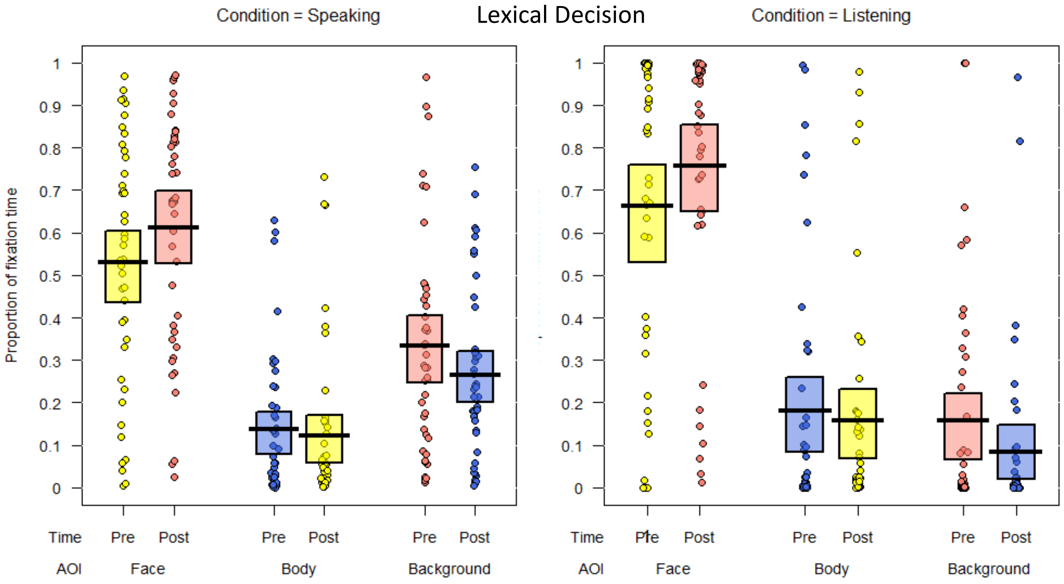


Figure 4. 9: The proportion of time spent fixating each AoI in each training group at the pre-and post-training session (averaged over age groups for illustration). The plots show raw data points, a horizontal line reflecting the condition mean, and a rectangle representing the Bayesian highest density interval.

4.3.6.2. Posters

Fixations to the three posters were analysed using a 3 x 4 x 2 x 3 x 2 mixed design ANOVA, crossing the between-subjects factors Age Group (adolescents *vs.* young adults *vs.* older adults) and Training Group (WM *vs.* IC *vs.* CF *vs.* LD) with the within-subjects factors Condition (speaking *vs.* listening), Time (pre *vs.* post training) and AoI (shared gaze *vs.* averted gaze *vs.* neutral). Full statistical effects are reported in Table 4. 12.

Table 4. 12: Statistical effects for fixations in the face to face conversation task. Asterisks show significance of effects, where * $p < .05$; ** $p < .01$; *** $p < .001$.

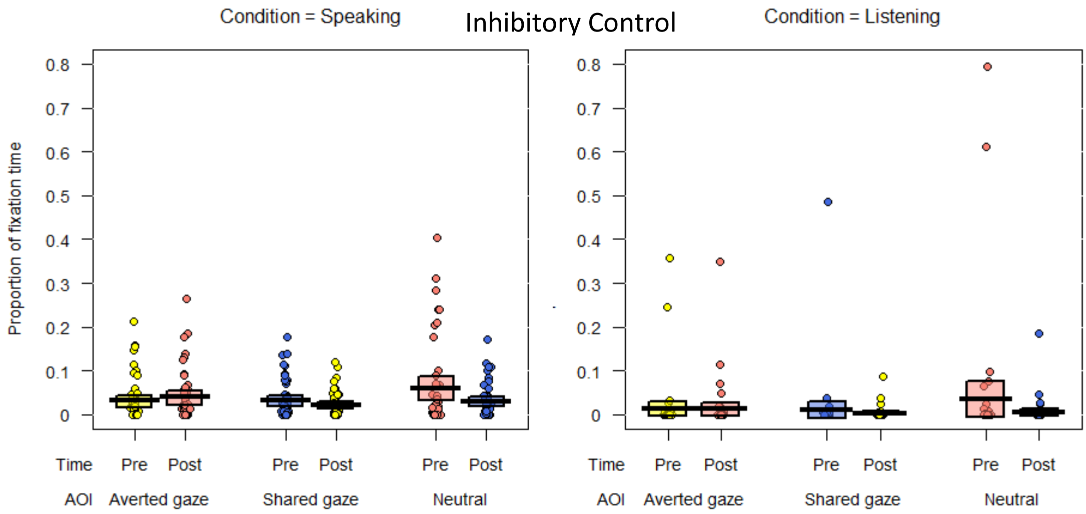
Effect	*df*	*F*	*p*	η_p^2
Age Group	(2, 179)	3.71	.03*	.04
Training Group	(3, 179)	2.07	.11	.03
Condition	(1, 179)	55.75	< .001***	.24
AoI	(2, 358)	1.45	.24	<.01
Time	(1, 179)	1.46	.23	<.01
Age Group x Training Group	(6, 179)	1.64	.14	.05

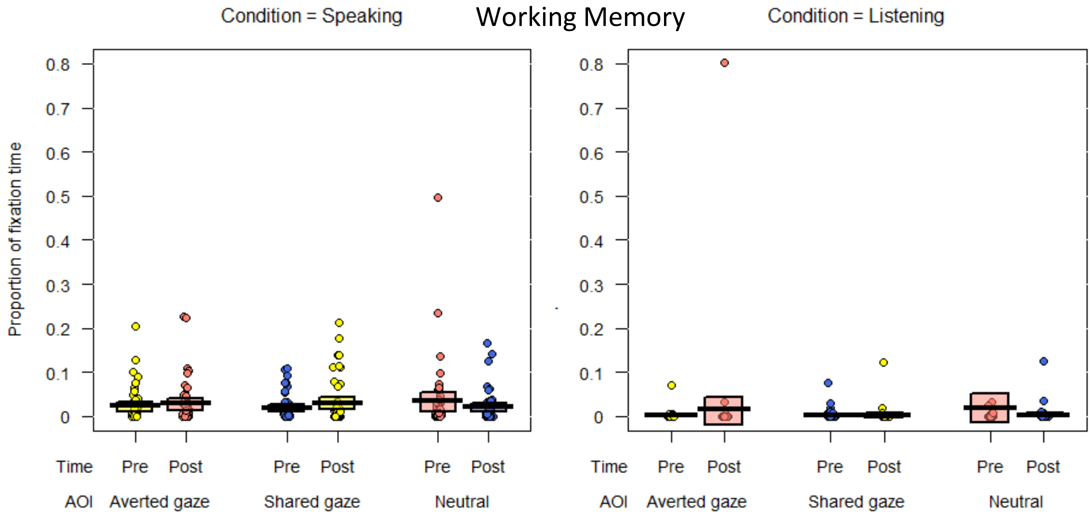
Age Group x AoI	(4, 358)	.17	.95	<.01
Training Group x AoI	(6, 358)	1.50	.18	.03
Age Group x Condition	(2, 179)	2.25	.11	.03
Training Group x Condition	(3, 179)	.86	.46	.01
Age Group x Time	(2, 179)	.36	.70	<.01
Training Group x Time	(3, 179)	.85	.47	.01
Condition x AoI	(2, 358)	.08	.93	<.01
AoI x Time	(2, 358)	4.67	.01*	.03
Condition x Time	(1, 179)	.50	.48	<.01
Age Group x Training Group x AoI	(12, 358)	.85	.60	.04
Age Group x Training Group x Time	(6, 179)	2.42	.03**	.08
Age Group x Training Group x Condition	(6, 179)	.45	.84	.11
Age Group x Condition x AoI	(4, 358)	1.12	.35	.01
Training Group x Condition x AoI	(6, 358)	.81	.56	.01
Age Group x AoI x Time	(4, 358)	.84	.50	<.01
Training Group x AoI x Time	(6, 358)	.76	.60	.01
Age Group x Condition x Time	(2, 179)	.87	.42	.01
Training Group x Condition x Time	(3, 179)	.48	.70	<.01
AoI x Condition x Time	(2, 358)	.17	.84	<.01
Age Group x Training Group x Condition x AoI	(12, 358)	1.32	.20	.04
Age Group x Training Group x AoI x Time	(12, 358)	.86	.58	.03
Age Group x Training Group x Condition x Time	(6, 179)	4.03	.007**	.12
Age Group x AoI x Condition x Time	(4, 358)	2.54	.04*	.03
Training Group x AoI x Condition x Time	(6, 358)	.91	.48	.02
Age Group x Training Group x AoI x Condition x Time	(12, 358)	.55	.88	.02

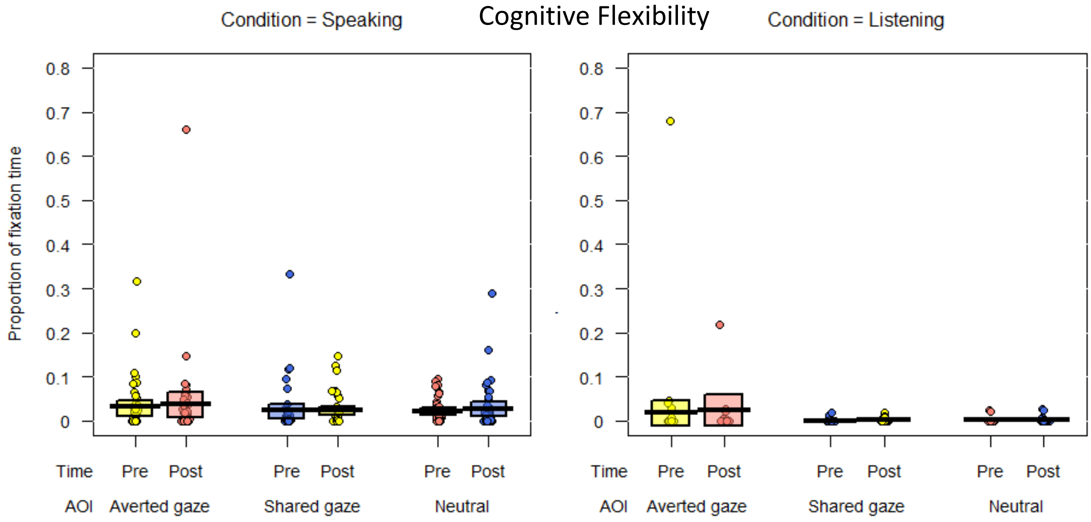
The main effects of Age Group and AoI replicated the patterns reported in Chapter 2. In addition, a significant AoI x Time interaction revealed that participants distributed their attention between the three poster AoIs differently at pre, $F(2, 1143) = 4.45$, $p = .01$, $\eta_p^2 < .01$, and post-training sessions, $F(2, 1143) = 4.31$, $p = .01$, $\eta_p^2 < .01$. Follow-up analyses showed that at the pre-training session participants spent a greater proportion of time looking at the neutral poster ($M = .03$) compared to the shared gaze poster (*M* = .02, *t*(190) = 2.06, *p* = .04), but looks to the neutral and averted gaze poster (*M* = .02) did not differ significantly (*t*(190) = 1.84, *p* = .07). At the post-training session, participants spent a greater proportion of time looking at the averted gaze poster (*M* = .03) compared to the neutral poster (*M* = .01, *t*(190) = 2.03, *p* = .04). No difference emerged comparing the proportion of fixations on the neutral vs the shared gaze poster (*t*(190) = .49, *p* = .63).

The Age Group x Training Group x Time interaction was significant, and was further subsumed under a significant 4-way Age Group x Training Group x Condition x Time interaction. Follow-up analyses showed that the Training Group x Age Group x Time interaction was significant in the listening condition, *F*(3, 179) = 4.28, *p* < .001, η_p^2 = .13, but not in the speaking condition, *F*(3, 179) = .61, *p* < .73. Examining effects separately for each age group revealed that the Training Group x Time interaction was significant in the older adult group, *F*(3, 56) = 2.81, *p* = .05, η_p^2 = .13, and adolescents *F*(3, 56) = 3.10, *p* = .04, η_p^2 = .14, but not in the young adult group, *F*(3, 67) = 2.23, *p* = .09. Specifically, adolescents in the control group looked longer at the posters in the pre- (*M* = .05) compared to the post-training session (*M* < .001, *t*(13) = 2.46, *p* = .03). None of the other comparisons from pre- to post-training reached significance (*t*s > .58, *p*s > .06).

The 4-way interaction Age group x AoI x Condition x Time was also significant and further explored running three separate 3 ways ANOVA on each AoI, however none of the ANOVAs revealed a significant interaction Age group x Time x Condition.


Condition = Speaking
Inhibitory Control
Condition = Listening
Proportion of fixation time
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Time
Pre
Post
AOI
Averted gaze
Shared gaze
Neutral


Condition = Speaking
Working Memory
Condition = Listening
Proportion of fixation time
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Time
Pre
Post
AOI
Averted gaze
Shared gaze
Neutral


Condition = Speaking
Cognitive Flexibility
Condition = Listening
Proportion of fixation time
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Time
Pre
Post
AOI
Averted gaze
Shared gaze
Neutral

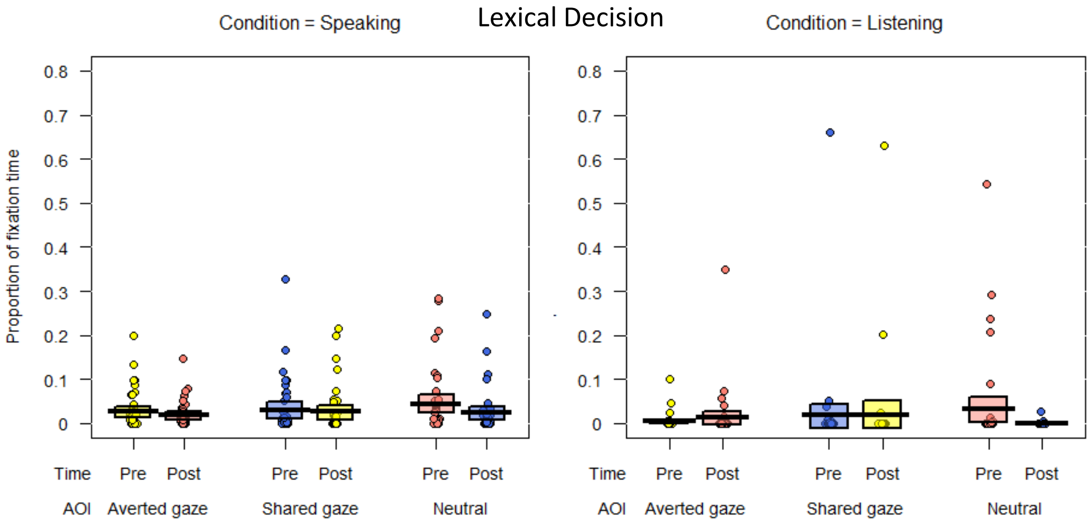


Figure 4.10: The proportion of time spent fixating each AoI in each training group at the pre and post-training session (averaged over age groups for illustration). The plots show raw data points, a horizontal line reflecting the condition mean, and a rectangle representing the Bayesian highest density interval.

4.3.7. Navigation in the real-world

The final sample in this task was 199 (63 adolescents, 73 young adults, 63 older adults); 29 participants were excluded due to insufficient eye-tracking data and 14 participants were excluded due to technical issues in one of their testing sessions. For this task, fixations were analysed using a 3 x 4 x 2 x 4 mixed design ANOVA, crossing the between-subjects factors Age Group (adolescents *vs.* young adults *vs.* older adults) and Training Group (WM *vs.* IC *vs.* CF *vs.* LD) with the within-subjects' factors, Time (pre *vs.* post training) and AoI (map *vs.* people *vs.* path *vs.* objects). Full statistical effects are reported in Table 4. 13[7].

---

[7] Note that some of the effects are not meaningful in this analysis because proportions of fixations for each participant/ time summed to 1.

Table 4.13: Statistical effects for fixations in the face to face conversation task. Asterisks show significance of effects, where * $p < .05$; *** $p < .001$.

	df	*F*	*p*	η_p^2
Age Group	(2, 187)	.02	.98	<.01
Training Group	(3, 187)	.07	.97	<.01
AoI	(3, 561)	1838.39	< .001***	.91
Time	(1, 187)	.01	.92	<.01
Age Group x Training Group	(6, 187)	.11	.99	<.01
Age Group x AoI	(6, 561)	3.85	< .001***	.04
Training Group x AoI	(9, 561)	.64	.76	.01
Age Group x Time	(2, 187)	.01	.99	<.01
Training Group x Time	(3, 187)	.03	.99	<.01
AoI x Time	(3, 561)	22.75	< .001***	.11
Age Group x Training Group x AoI	(18, 561)	1.87	.02*	.06
Age Group x Training Group x Time	(6, 187)	.05	1	<.01
Age Group x AoI x Time	(6, 561)	.54	.78	<.01
Training Group x AoI x Time	(9, 561)	.33	.96	<.01
Age Group x Training Group x AoI x Time	(18, 561)	.65	.86	.02

The main effects of AoI, and Age Group x AoI interaction replicated the patterns reported in Chapter 2. Importantly, the analysis revealed a significant AoI x Time interaction. Follow-up analyses compared fixations to each AoI between the pre- and post-training sessions. Results revealed that in the pre-training session participants looked more at the map ($M = .20$ *vs.* .14, $t(198) = 6.09$, $p < .001$) and less at the path ($M = .65$ *vs.* .71, $t(198) = 4.94$, $p < .001$) compared to the post-training session. Proportion of time fixating people or objects did not differ from pre- to post-training (*t*s < .46, *p*s > .65).

Finally, the AoI x Training Group x Age Group interaction was significant, however since this effect did not include the Time variable it most likely reflects basic group differences rather than training effects.

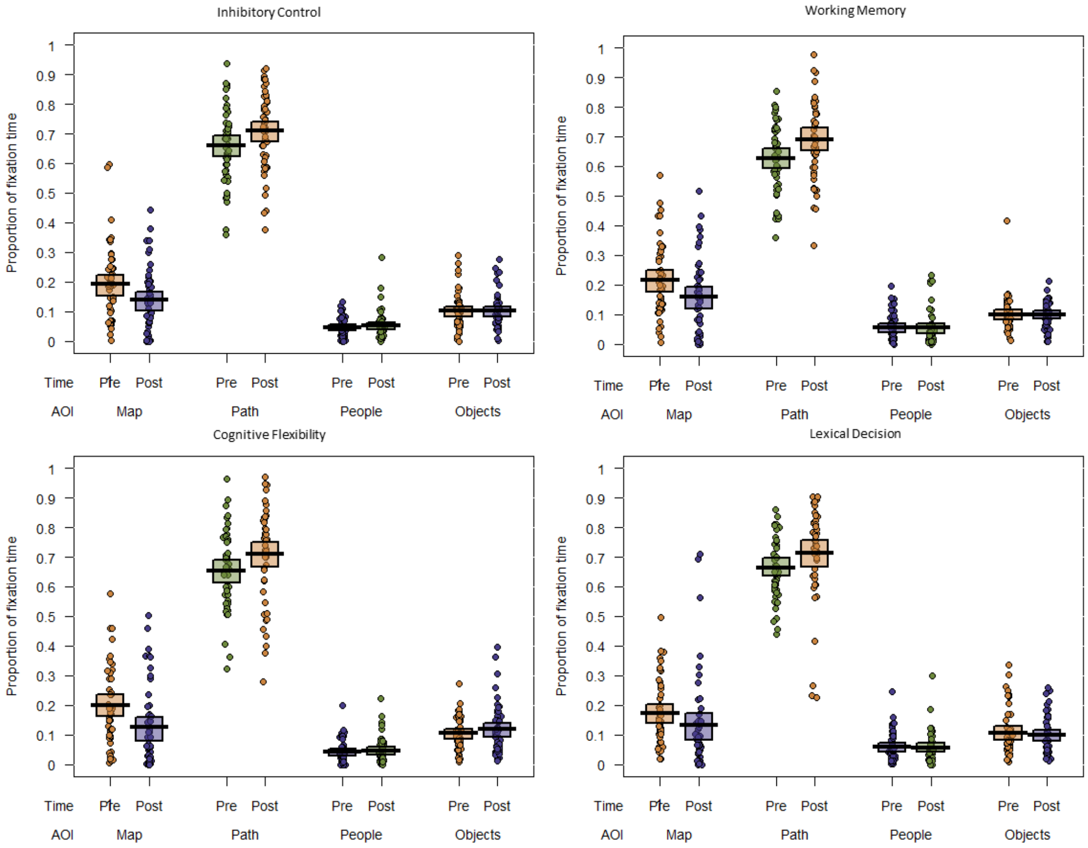


Figure 4. 11: The proportion of time spent fixating each AoI in each training group at the pre and post-training session (averaged over age groups for illustration). The plots show raw data points, a horizontal line reflecting the condition mean, and a rectangle representing the Bayesian highest density interval.

4.4. Discussion

In this Chapter, we tested whether cognitive and social abilities can be enhanced by training the underlying EFs, and whether these training effects differ in different age groups (adolescents, younger adults and older adults). First, we consider the degree to which the aging effects on various sub-components of social cognition reported in Chapter 2 were replicated here. As expected, more cognitively demanding tasks showed age-related differences. For example, age modulated the interaction between level and angle on the VSPT task, showing that level 2 judgements are more cognitively demanding than level 1. Similarly, in the Director task adolescents and older adults were less accurate than young adults to select the target object when there was a conflict between their own and the avatar's perspective (only in the pre-training session). Our measures of social attention also showed age-related decline, as older adults made fewer fixations towards their social partner compared to young adults during a face to face conversation, and in turn spent more time fixating the background (to manage the cognitive effort of maintaining the conversation). In the navigation task adolescents and older adults made fewer fixations towards people compared to young adults.

In contrast, other measures of social cognition were not subject to age-related decline, suggesting that they are relatively automatic and cognitively effortless. For example, older adults were more likely to make social inferences when watching videos of others interacting compared to young adults and adolescents, which suggests that people might become *more* sensitive to others' mental states across the lifespan. Finally, our empathy for others' pain task revealed that young and older adults showed a greater pain effect on explicit ratings and greater mu suppression for pain images than adolescents, suggesting that empathy brain networks develop then stabilise through the lifespan.

Regarding transfer effects between cognitive abilities, we expected to replicate our findings from Chapter 3, which showed limited indirect between cognitive skills. The analysis of indirect training effects across different EFs showed that participants' performance was enhanced on all tasks at the post-training session compared to the pre-training session, however this improvement did not differ across the four training groups. This replicates the pattern seen in Chapter 3 and reinforces our conclusion that improvements from time 1 to time 2 were due to practice with the tasks rather than indirect training effects that would lead to transfer of skills between specific tasks. As seen in previous research (Heinzel, et al., 2014; Karbach

& Kray, 2009; Reed, et al., 2014), performance on the cognitive assessment tasks differed between the three age groups, with young adults showing enhanced working memory, inhibitory control and cognitive flexibility compared to both adolescents and older adults.

Indeed, adolescents and older adults showed a greater improvement in cognitive flexibility from pre- to post-training compared to young adults, likely reflecting greater benefits of practice in these age groups who began with a lower baseline ability (Davidson, et al., 2006; Diamond, 2002). This age-related enhancement effect seen on our measure of cognitive flexibility can be interpreted in two ways. One is that these improvements reflect a mere practice effect, as seen on the other measures of EF. The second hypothesis links to previous literature that has revealed that WM (Gamboz, et al., 2009; Hartman, et al., 2001) and IC (Gamboz, et al., 2009) are both fundamental abilities to performance in the WCST, since WM allows us to maintain and manipulate incoming information from the sorting cards and IC helps us to inhibit a preponderant response in sorting cards according to a certain rule. Thus, it is possible that performance on the WCST improved from pre- to post-training because *all* EF training tasks tapped mechanisms involved in success for this task. It is important to note however, that performance also improved in the control group (i.e. trained in a lexical decision task, which did not target any EFs), which suggests that the former interpretation, based purely on practice effects, is more likely.

The main aim of this study was to explore a novel question on whether we can enhance social skills through EF training, thus showing indirect training of social cognition via its underlying cognitive mechanisms. We predicted that cognitive training could lead to improvements in those sub-components of social skills we found to be most cognitively demanding here and in Chapter 2 (i.e. level 2 perspective-taking, reference assignment, social attention), but would not alter performance on those sub-components of social abilities that are more automatic and less cognitive demanding (i.e. social inferencing, level 1 perspective-taking, empathy for others' pain). Overall, our findings revealed very limited evidence on the generalisability of training effects from cognitive to social domains following a 21-day adaptive training protocol.

Regarding the more cognitively demanding social tasks we found that performance improved overall from the pre- to post-training session, however the majority of these training effects were not specific to any training group. Therefore, similar to the findings for indirect training effects across different EFs, this pattern indicates that enhanced performance at time 2 is due to practice rather than genuine training of the underlying mechanisms. Specifically, on the VSPT task participants were faster overall to judge the avatar's perspective at the post-training session compared to the pre-training session, and at post-training, participants in the cognitive flexibility group responded significantly faster compared to the other three groups. This pattern might suggest that cognitive flexibility training taps the underlying skills for perspective-taking, such as shifting between own to other's perspective, which leads to better performance in this type of perspective-taking judgements. This interpretation would fit with previous studies that have shown a connection between cognitive flexibility and perspective-taking (Bradford, et al., 2015; Lin, et al., 2010; Long, et al., 2018). Importantly, however, none of our effects in this VSPT task involved condition variables, which makes it more likely that although training benefitted the overall speed of responding in this task (perhaps more so following the SST training task since it involved a similar response pattern), it did not target any of the specific mechanisms involved in perspective-taking. Moreover, this suggests that effects of practice impacted level 1 and level 2 perspective-taking in similar ways.

In the referential communication task, participants were more accurate and faster to respond at the post-training compared to pre-training session, and young adults only outperformed adolescents and older adults at the pre-training session, showing that age-differences in accuracy on this task were eliminated after practice. This suggests that practice has a greater enhancing effect on the ability to understand others when this ability is still under development (Dumontheil, et al., 2010) or subject to a decline (Hess, 2014). Importantly, since none of these changes over time differed by training group, we cannot attribute the enhanced responding on listener only trials to any specific enhanced cognitive mechanism. Further, our measures of real-world social attention (i.e. face to face conversation and navigation task) revealed that neither time or training group influenced the allocation of gaze towards social objects (i.e. the conversation partner’s face or people in the environment). This suggests that the way people of different ages distribute their attention in real-world settings is not affected by cognitive training, or even practice.

Finally, we predicted that tasks that tap more effortless sub-components of social cognition would not be affected by cognitive training; this prediction was confirmed in our data. Specifically, performance on level 1 VSPT was not modulated by the combined influence of time and training group, thus although participants responded faster when performing the task post-training this improvement was not tied to any specific EF, no did it differ from the control group. In the social hierarchy task, participants were overall more likely but slower to make social inferences at the pre- compared to post-training session (particularly beliefs and personality). Once again, however, these effects were not further modulated by an interaction with time and training group, suggesting that the effect of time reflects familiarization with the video stimuli across the two testing sessions. Further, in the empathy for others' pain task participants gave higher pain ratings to images at the pre- than post-training session, and this effect was not modulated by training group, which shows that participants may have become sensitised to painful images due to repeated exposure (Codispoti, et al., 2007; Preis, et al., 2015) but did not experience any training-related change in their pain empathy. Overall mu/alpha and beta suppression was greater at the post- than pre-training session, and this difference was greater over occipital than central electrodes, which reinforces our interpretation that cognitive training did not have a specific effect on empathy and instead suggests that the increase might reflect more general changes in participants' arousal between the two testing sessions.

None of the social cognition tasks revealed different training effects for the three different age groups, suggesting that indirect training following targeted training on EFs is not more likely in (healthy) groups who are not at their peak of social cognition abilities (i.e. adolescents and older adults).

4.5. Conclusions

To conclude, the EF training protocol failed to elicit consistent evidence for indirect transfer of trained cognitive abilities to any sub-component of social cognition. These findings raise important questions about the robustness of the relationship between social cognition and executive functions, and corroborate Chapter 3's finding of limited generalisability for skills learnt during cognitive training to different tasks that measure a related concept. As such, they identify limitations in this field, highlighting the specificity of effects to particular tasks or sub-components of EF and social cognition. We will discuss some of these limitations in Chapter 5.

In the next Chapter, I will summarise the novel findings reported throughout this thesis, and integrate them with existing empirical and theoretical accounts of social cognition to address the three key questions I set out at the outset of this thesis. Namely, does social cognition differ across the lifespan? Can we train EFs? Can we train social cognition indirectly through cognitive training? I will also reflect on some of the key limitations and future directions for this work.

CHAPTER 5: GENERAL DISCUSSION

5.1. Overview

The broad aim of this thesis was to investigate whether and how social cognitive abilities change across the lifespan, whether they can be enhanced through cognitive training, and whether these training effects are more effective in different age groups (i.e. when social and cognitive skills are still developing or in decline).

In Chapter 1, I provided a definition of social cognition, which includes perspective-taking, ToM and empathy, and outlined some of the key theories that have tried to explain its functioning. Extensive research has focused on certain aspects of social cognition, such as how we interact in the social environment, how we try to interpret and understand other’s mental states, and how we empathise with others. Previous studies in this field have mainly focused on one aspect of social cognition (e.g. perspective-taking), adopting a specific measure (e.g. false belief task), usually in single a targeted population (e.g. children). I highlighted limitations of this approach, emphasising that focusing on just one isolated sub-component of social cognition does capture the complexity of human behaviour and the variety of our social interactions. Moreover, investigating specific abilities in targeted populations has led to mixed evidence on whether social cognition varies across the lifespan. Therefore, to address these limitations, in Chapter 2 I presented research that examines social cognitive abilities comparing adolescents, young and old adults’ performance in a battery of tasks that employed explicit and implicit measures to assess distinct sub-components of social cognition. Specifically, to obtain a more comprehensive understanding on whether social cognitive abilities differ between these three age groups, we adopted six different tasks alongside behavioural responses (i.e. explicit measures), brain responses detected with EEG, and eye movements in the real world recorded with mobile eye-tracking (i.e. implicit measures).

In Chapters 1 and 2 I reported evidence that social cognitive abilities are impaired in aging and clinical conditions, highlighting attempts from previous research to improve these abilities through training. However, the efficacy of this type of training interventions is still unclear. I also reported how ToM and cognitive functions are related, detailing evidence from the neuroimaging literature that has revealed overlapping brain activity during social and cognitive tasks. In support of this correlation, aging studies have reported that social and cognitive skills seem to moderate each other across the lifespan, however whether social sustain cognitive abilities or vice versa, and if they follow the same developmental trajectory across the lifespan (i.e. increment until the early adulthood and a decline in the old age) remains under debate. Among the theories that have tried to explain the relationship between the social and the cognitive components, I have referred to the ‘two systems model for mindreading’, proposed by Apperly and Butterfill (2009, 2011) to explain a distinction between effortful and automatic processes involved in ToM functioning. This model helps to delineate how certain social abilities are subject to changes and/or to decline, whereas others are intact, or even enhanced in old age. Empirical research has shown there is a general agreement that EFs are a good predictor for social cognition (Carlson et al. 2004; Razza & Blair, 2009). However, only a few studies have tried to improve social cognition by training these underlying EFs, and these studies have been limited by training only a single EF, assessing a sub-component of social cognition on a single measure, and targeting a specific age group or clinical population. As such, the efficacy of this kind of intervention has not yet been established and further investigations are needed.

Based on this work, in Chapter 3 I developed a 21-day adaptive cognitive training programme that specifically targets Working Memory, Inhibitory Control or Cognitive Flexibility, with the aim to assess the generalisability of cognitive improvement from training to other tasks that measure the same EF and tasks that measure related EFs. Finally, in Chapter 4 I applied this 21-day adaptive cognitive training programme to assess whether different sub-components of social cognition can be enhanced through cognitive training, and whether effects differ in different age groups. In the next sections I will discuss the results relative to the three research questions I aimed to address, detailing some of the limitations we found in this work.

5.2. Does social cognition differ across the lifespan?

The aging literature has reported diverging evidence on how social cognition, in particular perspective-taking, empathy and social attention, differs across the lifespan.

In Chapter 2, we found mixed results on whether social cognition is impaired in adolescents and older adults compared to young adults. Overall, response times on various tasks were slower in older adults compared to the two younger groups; in the VSPT task older adults were slower to judge *where* an object was placed relative to an avatar or *what* an avatar could or could not see on an object, in the referential communication task older adults were slower to select the target object, and in the hierarchy of social inferences task older adults were slower to make social inferences. These results are likely to reflect a general cognitive slowing that has been shown in many previous aging investigations (Salthouse et al. 2000; Verhaeghen, 2011). More interesting, however, some tasks showed specific impairments in social cognition among adolescents and older adults compared to young adults. The ability to make level 2 judgements about how the avatar could see the cube/number in the VSPT task was subject to age related change, with young adults showing steeper increments in reaction times with increasing angles; level 1 judgements did not differ with age. When using perspective to interpret reference in the Director task, young adults outperformed both adolescents and older adults when they held privileged knowledge about a competitor object (i.e. listener only condition), and only older adults were delayed in selecting the target object in the listener only than shared view condition. Finally, our two measures of social attention in the real-world revealed reduced attention on social stimuli among adolescents and older adults compared to young adults. Specifically, during the face-to-face conversation task we found that adolescents and older adults made fewer fixations on the experimenter's face and looked longer at the background. In the navigation task, adolescents and older adults spent less time looking at people in the environment. These results might suggest a difficulty to allocate attention on complex and dynamic social stimuli during cognitively demanding tasks, such as a real-life conversation and navigation in real world. Overall, evidence from these four tasks of social cognition suggest that age-related developments and declines occur across multiple sub-components of social cognition, and are likely to represent broad changes in mentalizing ability. These patterns therefore fit with the predictions from Apperly and Butterfill (2009)'s model, that some sub-components of social cognition are cognitively demanding and therefore subject to a decline in older age.

However, we also found evidence that some sub-components of social cognition are stable or even enhanced over the lifespan. Older participants were more likely to infer mental states when watching others interacting in the hierarchy task compared to both adolescents and young adults. The older participants therefore seem to show a greater sensitivity in detecting mental states, despite showing explicit difficulties in using these mental states to predict others' actions in more cognitively demanding tasks (e.g. Ferguson, et al. 2015). It is important to note however, as mentioned in Chapter 2, that older adults' performance in this task might reflect an over use of cause-effect inferences. In the empathy for others' pain task, behavioural ratings revealed a greater pain effect (pain *vs.* no-pain images) among young and older adults compared to adolescents. EEG data revealed that older adults showed overall lower mu/alpha and beta suppressions compared to young adults, and these age differences were greatest over the occipital sites. This general pattern is likely to reflect enhanced attentional processes among young adults (Hobson & Bishop, 2017; Perry et al., 2010). More importantly, the three age groups differed in their mu (alpha and beta) suppression response to pain. In the alpha range, mu suppression to pain images was greater in the young and older adults compared to adolescents, and in the beta range mu suppression to pain images was greater among older adults compared to both adolescents and young adults. These findings demonstrate that mu/alpha rhythm can reveal age-related changes in empathy (Isaacowitz & Stanley, 2011; Ruffman, et al. 2008; Sullivan, et al. 2017), but importantly that older adults are not impaired at recognising and responding to others' pain. These patterns are also consistent with the predictions from Apperly and Butterfill (2009)'s model, that some sub-components of social cognition are relatively automatic and cognitively efficient so are less susceptible to age declines.

The distinct patterns of social impairments with age that we have revealed in different sub-components of social cognition also inform theoretical models. As described in Chapter 1, Apperly and Butterfill (2009, 2011; see also Apperly, 2009) detailed one of the first theoretical frameworks that links social abilities with cognitive mechanisms, the ‘two systems model for mindreading’. Applied to lifespan development, this model proposes that basic social inferences that do not rely on cognitive abilities are spontaneously activated much earlier than four years old, the age at which children are known to pass explicit tests of false beliefs. The model also proposes that development of more sophisticated forms of social interaction continues through childhood and adolescence as they rely on increasingly complex cognitive mechanisms, which are known to develop over a protracted period into early adulthood (Best & Miller, 2010; De Luca et al., 2003). Finally, the model accounts for a decline in more cognitively demanding social abilities into older age, as age-related declines in cognitive functioning are relatively robust due to changes in the frontal lobes, specifically age-related volume reduction in the prefrontal cortex (Gunning-Dixon & Raz, 2003). Carruthers (2016) suggested a similar model linking social and cognitive processes, but argues that a single mindreading system can account for the need to recruit cognitive resources during some mindreading tasks. This single system operates in a fairly rudimentary way in early infancy, based on a set of conceptual primitives and thought-attributions. It becomes increasingly efficient from infancy to childhood through a continuous period of development, as social and communicative experience grows, cognitive and language mechanisms mature, and the connection between mindreading and cognition strengthens. Importantly, this model predicts that success in social interaction will vary depending on the demands placed on executive function and language, both of which are subject to age-related decline.

Our data provide support for this link between social and cognitive processes and seem most consistent with the predictions of the ‘two system model for mindreading’ proposed by Apperly et al. (2009, 2010). As predicted by this model, those aspects of social cognition that are cognitively efficient did not show changes with reduced cognitive abilities, whereas those cognitively demanding aspects of social cognition declined with age-related reductions in cognitive abilities. At the same time, our data could also be explained by the single mindreading model proposed by Carruthers (2017), which suggests that EFs can be recruited when mentalising when required. However, in contrast to the Apperly model, this model would suggest that any ToM task could be cognitively effortful if EFs are weak (e.g. in infancy), which is at odds with our finding of some seemingly effortless ToM inferences even in older age when EFs have deteriorated. Importantly, both models can account for the different scenarios we can encounter in our everyday life and reflect the complexity of social and cognitive processes across the lifespan. Further research is required to systematically manipulate cognitive and language constraints on social cognition to distinguish these two models.

How we can explain these aging differences? The answer might be found in structural changes in the brain areas that are involved in social cognition. Major changes can be found in prefrontal areas that undergo synaptic pruning during adolescence and in older age; grey matter reduces and functional connectivity in fronto-parietal neural networks decreases (Madden et al. 2010), which leads to a consequent recruitment of additional brain areas (Hong et al., 2016). Moreover, changes in white-matter and impoverishment of brain volume (Rabbitt et al. 2007a and 2007b) seem to underly a general slowing in older adults. It therefore seems likely that these changes in brain structure and activity can sustain differences in performance in adolescents and older adults compared to young adults.

The current empirical work aimed to address some of the key limitations form previous work. For instance, the mixed evidence on aging and social cognition could reflect the use of different tasks to measure different sub-components of social cognition (e.g. Strange stories or False beliefs tasks) via diverse modalities (e.g. audio, visual). Therefore, divergent performance on ToM across the lifespan can be a result of adopting a single specific measure presented through a certain modality rather than a general aging effect. Another limitation of previous work is that explicit behavioural measures are not sufficient to assess the great variety of humans' social cognition. Hence it is important to adopt a large battery that employs different stimuli and examines social abilities through diverse modalities. In the current study we also addressed previous limitations of small samples, usually focused on one age group and one aspect of social cognition. In fact, our results show the importance to adopt a large battery of tasks that tap different sub-components of social cognition and to use a broad range of explicit and implicit measures. Nevertheless, though our general approach, testing multiple sub-components of social cognition, using a range of measures, in large samples of participants that vary across the lifespan, provides valuable new insights to this area, some limitations remain.

First, earlier studies have demonstrated a difficulty in allocating attentional resources in older age. In particular, older adults seem to struggle in focusing on certain aspects of a scene, certain stimuli: in fact, older adults experience difficulty in suppressing irrelevant distractors (Gazzaley et al., 2005; Zanto, et al. 2010) and employing selective attention (Engle & Kane, 2004). These more general attention impairments in older adults might therefore underlie some of the difficulties that emerged in our social cognition tasks, since all those tasks that revealed impairments in older age required participants to focus their attention on specific features of a scene and ignore distractors. Future investigations might answer this key question by including measures of cognitive attention, and using this in statistical models to partial out influence from general attention and isolate the social attention processes more specifically.

Second, in the eye-tracking tasks it is possible that some age differences might reflect non-social differences between age groups. For example, older people might be less confident in walking due to reduced motor capacities or problems related to their posture, which in turn forces their attention away from the social environment and onto the path and obstacles in front of them. Older adults might also have reduced vision, meaning that they fixate longer on fixed physical objects that can help them to navigate in an unfamiliar space. Finally, older adults experience greater difficulties with memory and as such may be more reliant on frequent checking of the map and signage to stay on task.

Another important point to consider is the use of implicit measures to assess age-related changes, such as brain activity and eye-tracking fixations. The quality of data detected with EEG and eye-tracking can change between different ages. As highlighted by Zappasodi et al. (2015), “progressive neural specialization and global integration of the brain networks during development and maturation, as well as the loss of synaptic connections and neuronal apoptosis in physiological brain aging, also result in a change of dynamics of the electrophysiological data”. Similarly, eye-tracking literature has pointed out how eye movements detected in the lab can be affected by aging (Dowiasch, et al. 2015; Spooner, et al. 1980). Moreover, the use of implicit measures without a direct explicit measure means that we must infer the social/cognitive processes or behaviours that the implicit effect reflects, and this might not always be accurate. It is relevant to keep in mind these considerations when looking at age differences across the lifespan.

5.3. Can we train EFs?

Broadly, our research on cognitive training addressed some important concerns about training interventions raised from previous studies, and aimed to provide a more thorough understanding of the generalisability of training effects within and between EFs. Firstly, we employed an appropriate control condition (i.e. an adaptive training protocol that did not target any EF) in order to compare potential training effects in the experimental groups. We specifically avoided using a passive control group with limited contact with the experimenters (i.e. only during the pre- and post-assessment sessions), or an active control group that practised with an activity that is not cognitively demanding. Another important point was to challenge our participants with an adaptive training. Practicing with no changes in difficulty can prevent participants from pushing themselves, so that might lead to no improvements. Lastly, we recruited a large sample size to identify and compare training effects; we tested 160 participants for the experiment reported in Chapter 3, and 232 participants for the experiment reported in Chapter 4.

Results from Chapter 3 are in line with previous findings, reporting robust evidence of direct transfers (improvements in the same trained task), however limited indirect transfers (near and far) were detected. These findings contrast with previous studies that have shown near transfer improvements following WM training (Heinzel et al., 2014, 2016; Maraver et al., 2016; Thorell et al., 2009; Li et al., 2010), IC training (Berkman et al., 2014; Enge et al., 2014; Thorell et al., 2009), or CF training (Karbach & Kray, 2009). In Chapter 3 we tested younger adults who are at the peak of their cognitive functioning. This might represent a limitation in finding transfers to other tasks due to an already close-to-ceiling performance. In Chapter 4, we expanded our sample to include adolescents and older adults as well as young adults, but again we found evidence for indirect training across different tasks that measure the same EF, or across tasks that measure different EFs. Both chapters showed improvements with practice from pre- to post-training, but no evidence for genuine enhancement in cognitive functioning. This practice effect was particularly noticeable on the WCST measure of task switching as adolescents and older adults performed significantly higher at the post-training session, likely due to a lower baseline performance compared to young adults and therefore greater opportunity to benefit from practice. In sum, we can then conclude that overall, the efficacy of cognitive training interventions is inconclusive, and more investigations are needed.

Based on previous research, we might have expected to see training based improvements in adolescents and older adults since the brain is undergoing more structural changes. In their metanalysis, Karback and Verhaeghen, (2014) for example reported clear near but limited far-transfers in old age after training in WM or CF. Zinke et al. (2012) found a reduction in switching costs in adolescents (10 – 14 years old) after CF training, and reported near transfer effects to other CF measures, far transfer to WM tasks and a general reduction of reaction times. Our results did not confirm any clear indirect training effects.

When interpreting the current findings, it is important to note that in general, transfers are expected when there is a high degree of overlap between the underlying processes involved in the different tasks (Buschkuehl, et al. 2012). In this work, we have included different assessment tasks from those adopted in the training, so that any transfers could be related to improvements in performance in that cognitive ability rather than shared strategies or response requirements between tasks. It may be then, that near transfer training effects did not emerge because the tasks we chose did not share sufficient cognitive processes to those used during training. Relatedly, issues around task impurity (i.e. the degree to which multiple EFs are recruited in different cognitive measures) mean that it is difficult to accurately isolate just one EF for training or assessment, and performance is likely to reflect contributions from several EF components. Previous research has shown that EF processes moderately correlate with each other (i.e. unity) but are also distinct (i.e. diversity) from each other (Miyake et al. 2000; Friedman et al. 2006).

There are also possible limitations with the choice of dependent variables we used in each task. To keep our analyses consistent across the training tasks we analysed accuracy or reaction times. However, other measures could be explored such as variance in the Stop Signal Reaction Times (SSRT) and flanker effect (rt compatible trails – rt incompatible trials), or absolute scores for the OSpan (i.e. number of letters recalled in the right order). It is possible that implementing different dependent variables would yield different results since they might tap a slightly different construct. Future work should aim to balance the need for consistency across tasks, and optimal task measures.

Another possible limitation to detecting training effects, is that our statistical analyses did not control for participants' performance at baseline, however how participants scored at the pre-training session is likely to predict their performance at the post-training session and their capacity to improve. The literature proposes two accounts to explain this: the magnification account in which participants with high performance before the training intervention will obtain higher scores at the post-training session, and compensation effects in which participants who gained the lowest score at the pre-test, achieve higher scores at post-training session (Lövdén, et al. 2012). Follow-up analyses therefore should aim to include baseline performance as a covariate.

In contrast with some other studies that adopted a mixed approach to train multiple EFs in one protocol (Nouchi et al. 2012; Ballesteros et al. 2014) or an intervention that targeted one specific EF (Jaeggy et al. 2008), our training focused on training several individual components of EFs. This seems to be a strength because training multiple EFs at the same time might reduce the time spent training a specific cognitive component, diminishing the power of transfer effects. Moreover, limiting the training protocol to just one component of EF would not explain the variety of difficulties or strengths that can be encountered in a sample, especially in diverse aging populations. Further analyses might aim to examine whether specific individual difference features help us identify those individuals who might benefit most from a cognitive training protocol. For instance, it has been suggested that older age individuals are likely to vary more from each other than young adults do (Buitenweg et al. 2012).

In sum, the EF training interventions reported here contribute to debates around the efficacy of training interventions across the lifespan, and the key methodological considerations. It is notable that the training literature on adolescents is very small. For instance, WM training with adolescence is restricted to early adolescence (~ 11 years old; Loosli et al. 2012). Cognitive interventions with older adults are typically restricted to WM and CF, whereas less attention has been given to training IC. Previous research has used IC measures as an assessment task (Davidson et al. 2003), or has adopted the Stroop task for IC training but not included a control group for comparison (Wilkinson & Yang, 2012). Further research is needed to further explore this field, particularly to understand the conditions under which cognitive skills transfer, and how cognitive training effects are influenced by neuroplasticity and physiological changes across the lifespan (Park & Bischof, 2013; Ballesteros et al. 2018).

5.4. Can we train social cognition through cognitive training?

Our results showed limited generalisation effects of cognitive training interventions on social cognition. Following previous studies highlighting the relationship between EFs and perspective-taking (Hartwright et al. 2012; Lin, et al. 2010, Qureshi & Monk, 2018; Bradford et al. 2015; Cane et al. 2017; Surtees et al. 2013b) we expected to find improvements in those sub-components of social skills that are most cognitively demanding. First, our data replicated the lack of indirect training across different tasks of EFs, or across different components of EF. More importantly, while most of our social cognition tasks showed a general improvement in performance from the pre- to post-training session, these effects were not specific to any training group. This suggests that general improvements on the tasks reflected enhanced familiarity with the tasks, and that our 21-day cognitive training protocols did not sufficiently alter the cognitive processes that underlie social cognition.

Our results are therefore in line with some earlier investigations that have failed to observe generalisability of far transfers to other far constructs, such as academic skills and emotion regulation (Melby-Lervag & Hulme, 2013; Sala & Gobet, 2017). A crucial aspect of effective training protocols seems to be the extent to which the trained and the assessed abilities share similar procedures, therefore future research should aim to employ tasks with greater overlap in these procedures. In sum, despite the use of a big sample and a wide battery of tasks, our results suggest that training unique EFs cannot enhance transfer effects on more distant and complex abilities.

These findings are in contrast with improvements in social cognition that have been reported by previous research (Kloo & Perner, 2003; Santiesteban et al., 2012). It is likely that methodological differences between the studies are responsible. For example, Kloo and Perner found improvements in false belief understanding following a cognitive flexibility training in 3 and 4 year old children. The authors adopted a Dimensional Change Card Sorting task (i.e. DCCS) as a measure of cognitive flexibility, which has previously been found to be correlated with the false-belief task (Carlson & Moses, 2001; Perner & Lang, 2002). As such, it is likely that the EF training and social cognition assessment tasks Kloo and Perner used had a higher degree of procedural overlap than the tasks employed here. In addition, the population tested was in an age period that is particularly sensitive to ToM changes (3-4 years old), and might therefore be more likely to show improvements from a baseline as a result of training. In another study, Santiesteban et al. (2012) revealed that healthy young adults showed improvements in perspective-taking following training in imitation-inhibitory control, but not after either imitation or inhibitory control training. This suggests that training either social or cognitive abilities in isolation, cannot enhance ToM, and that more effective transfer of skills can be achieved by training combined aspects of social cognition (e.g. Bradford et al. (2015) showed that adults' false belief performance correlated with a social face-Stroop task). Moreover, transfers were limited to perspective-taking (i.e. the ability to distinguish between the self and others), but did not generalise to other measures of ToM such as false belief understanding. It is also important to acknowledge that Santiesteban et al. (2012) did not include a baseline assessment of ToM performance before the training intervention, which makes it difficult to quantify the real efficacy of the training intervention.

An important limitation of our work is that we did not assess whether age-related changes on our social cognitive tasks were genuinely mediated by changes in EFs, and if so which ones. This link was assumed based on previous research showing clear development and declines through adolescence and older age, respectively, but due to time and space limitations in the PhD we did not formally test it. Further analysis of the current data is needed to test whether each of the sub-components of social cognition tested here relied on EFs to better understand the links between the EF training tasks and social cognition. Moreover, since our results from Chapters 3 and 4 failed to find any evidence of indirect training from one EF task to another EF task, we can assume that the underlying EFs have not been trained effectively, thus the chance of additional generalisation to the social domain is weak.

5.5. Methodological considerations

Although the use of ecologically valid methods (i.e. real-world eye-tracking) is a key strength of the current work, employing such unconstrained methods also raises some limitations. For example, people might alter their real-world looking behaviour while wearing eye-tracking glasses, as knowing that their gaze is being monitored makes them feel more self-conscious about where they are looking. The effect of being watched is likely to have a particular impact in looks to social content in our tasks, as participants avoid staring at other people to conform with cultural norms and as a means of reputation management. In line with this, Canigueral and Hamilton (2019) revealed an ‘audience effect’ whereby people were less likely to look at their conversation partner, but more likely to act pro-socially, when they believed they were being watched live by the other person. In addition, we note that the number of people that participants encountered during the navigation task, and the context in which they appeared (e.g. individuals *vs.* groups), differed between participants due to factors such as time of day, stage of the academic year etc. Since participants were tested throughout a two year period and age groups were recruited in parallel, we do not think these differences could have influenced age effects on social attention, however it is important to note that lack of availability may have contributed to the overall low proportion of time spent fixating people in this task. Foulsham et al. (2011) recorded social attention in an outdoor University setting and reported ~22% of gazes were on people, compared to ~5% in the indoor University setting here. Nevertheless, both tasks showed the same general pattern that people made fewer fixations to people compared to either objects or the path.

More generally, it is possible that characteristics of the people in our real-world environments elicited in/out-group effects on social attention (Simpson & Todd, 2017; Todd, et al. 2011; Savitsky, et al. 2011), and that these biases may have influenced age group effects. Specifically, the experimenters who led the face-to-face conversation were young adult females (aged ~25 years old), and the majority of people encountered in the navigation task were young adult students due to the campus University setting. Previous research has shown that an own-age bias enhances performance in a range of social perception tasks (e.g. Bailey et al. 2014; Melinder, et al. 2010; Rhodes & Anastasi, 2012; Slessor, et al. 2014), including heightened attention towards faces that are in the same age category as the perceiver (e.g. Bailey et al. 2014), superior memory for faces of one's own age group (Rhodes & Anastasi, 2012), and enhanced eye-gaze following for own-age faces (Ferguson, et al., 2018; Slessor et al. 2010). Indeed, young adults may be more susceptible to these in-group biases than older adults (Slessor et al. 2010). Thus, future research is needed to systematically vary the social context to explore whether young adults' increased likelihood to fixate on the experimenter’s face during the face-to-face conversation or to people in the navigation task reflects a general social processing advantage in this young adult group, or a more specific preference to look at people from one’s own age group. Moreover, it will be important to understand whether the reduced social attention seen here among adolescents and older adults is attenuated when more own-age people are available in the environment. This is an especially interesting question for adolescents who are particularly sensitive to their social environment (Peper & Dahl, 2013), and are thought to preferentially orient to their same-aged peers (Blakemore & Mills, 2014). Understanding the factors that influence real-world social attention is a vital next step.

Another important methodological consideration stemming from the work presented here relates to the choice of training intervention. As discussed in Chapters 1, 3 and 4, researchers have tested numerous interventions with the aim of enhancing peoples’ cognitive capacities and social wellbeing. An exciting new avenue of research on the ‘social brain’ has involved the use of Transcranial Direct Current Stimulation (tDCS), a safe, non-invasive technique for modulating neural activity by applying a weak current to the skull (Sellaro et al., 2016; Santiesteban et al. 2015; Adenzato et al. 2017; Martin et al. 2017). Martin et al. (2017) state that excitatory “anodal” tDCS *increases* the likelihood of neuronal firing, whilst inhibitory “cathodal” stimulation *reduces* the likelihood of neuronal firing. Studies using tDCS have reported intriguing results, including the presence of sex differences in the results outcome; for instance, tDCS administered to the dorsal-medial prefrontal cortex (dmPFC) showed improved cognitive ToM performance for females only, with no improvement in male participants’ performance, in emotion recognition (Martin et al., 2017) and attribution of intentions (Adenzato et al., 2017) tasks. It has also been shown that tDCS administration can enhance adaptive cognitive control in both younger (Gbadeyan et al., 2016) and older (Gbadeyan et al., 2019) adults, highlighting the potential beneficial effects of tDCS on cognitive abilities. These results provide a promising basis for future research, including examining improvements in social cognition task performance that may be encouraged by tDCS administration for individuals at different ages experiencing declines in their social cognition capacities.

Another approach to cognitive decline intervention is increased engagement in physical activity, particularly aerobic and strength exercise, which have been shown to benefit cognition abilities (e.g., Colcombe et al. 2006; Erickson et al. 2011; Voss et al., 2010). These benefits are argued to be selective, with studies showing that increased physical activity specifically leads to an improvement in executive control processes (Colcombe & Kramer, 2003; Kramer & Willis, 2002). These results suggest that physical activity can positively impact cognitive functioning in older age. Given that research has indicated a strong link between social cognition abilities and executive functioning capacities (e.g., Bailey & Henry, 2008; Cane et al., 2017; Bradford et al., 2015; German & Hehman, 2006), improvements in executive functioning as a result of engagement in physical activity could ultimately also support improvements in social cognition abilities in older age. Gheysen et al. (2018) suggest that the optimum intervention for mediating cognitive declines seen in older age may be to combine *both* physical activity and cognitive interventions, with the key component of interventions being that they challenge the individual to allow positive cognitive effects to be achieved (see also Zhu et al. 2016).

5.6. Future directions

In terms of potential training effects, future research should aim to probe a broader range of transfer effects, using different measures to assess quality of everyday life pre- and post-training, with particular attention to social and cognitive skills. Anecdotally, some of our participants, especially in the older adult group, reported some improvements in everyday life such as remembering more information, and sustaining attention. This reinforces the importance of assessing participants' real-world experiences of social and cognitive functioning, alongside more controlled lab-based assessments.

Due to constraints of the PhD, we did not have the time to test a middle-aged adult group (i.e. participants aged 40-60), leaving open questions about when age-related changes first emerge in the different sub-components of social cognition and EFs. As mentioned in Chapter 1, age-related changes in EFs are well established, but differences in their magnitude and trajectory exist between different sub-components of EF. Some aspects of cognitive decline begin from 20-30 years old, and decreasing at a faster rate with increasing age (Salthouse, 2009; Singh-Manoux et al., 2012). However, Nilsson et al. (2009) pointed that these claims have been based on cross-sectional studies, whereas when looking at longitudinal data, the cognitive decline starts much later (i.e. ~ 60 years). Recently, some studies have begun to examine age-related changes in social cognition in middle age. For example, Brunsdon et al. (2019) found that mirror neuron activity during action observation increases linearly throughout adulthood and not from the onset of older age, around 65, as previously thought. Similarly, Bradford et al. (2020) examined the brain's N400 response to narratives and found that adults became increasingly likely to interpret false-belief events from an egocentric perspective with increasing age (linearly from 10-86 years of age). More research is needed to further explore social cognitive processing in middle aged adults.

Although we employed a battery of tasks to measure social and cognitive functioning, it would be useful to continue expanding this battery, or conducting secondary analysis of additional measures for the existing data, to assess specific predictions about social cognition across the lifespan. For instance, in the empathy for others' pain task it would be worthwhile to supplement the mu/alpha suppressions analysis with ERPs, as in Cheng et al. (2014) and Mella et al. (2012). Studying ERPs components provides a valuable insight to the temporal nature of brain responses, and can be indicative of emotional and cognitive arousal in response to empathy stimuli, therefore could help to identify an attentional and an empathic response. Future data collection could also aim to target specific sub-components of social cognition and systematically increase the cognitive effort involved to test predictions about the role of EFs (i.e. test the predictions from Apperly and Butterfill, 2009) and impact on aging populations. As mentioned in Chapter 1, social cognition is characterized by different sub-components and each of these can be assessed through different tasks. Given the great attention that has been given to mindreading, future research might also include other tasks focused on emotion recognition, which seems to differ across genders (Martin et al., 2017; Richter et al., 2011), age (Mill et al., 2009; Philips et al., 2002; Richter et al., 2010; Sullivan & Ruffman, 2004b) and psychiatric disorders (Brüne, 2005b). In addition, researchers should aim for ecologically valid measures alongside tightly-controlled lab-based tasks, for example assessing everyday life quality, or using virtual reality to immerse participants in social situations in which individual variables, such as a virtual character’s eye gaze or in/out-group membership, can be systematically manipulated while other factors are fully controlled.

Finally, our sample was taken from a healthy population, who were confident in using a PC, and had no particular health limitations, or current neurological or psychiatric disorders. This suggests an already high level of general functioning among our participants and consequently they may have been close to a ceiling effect on performance. Future research might wish to investigate whether the training intervention is more effective with clinical populations, as suggested by previous studies with autistic (Fisher & Happé, 2005; Golan & Baron-Cohen, 2006) and schizophrenic participants (Hooker et al., 2012).

5.7. Conclusions

Understanding others involves a complex set of social and cognitive skills. It requires us to understand and predict others’ behaviour, by inferring their mental states and emotions, but also distinguishing these mental states from our own. Inferring others’ intentions, beliefs, and desires relies on information from different sources, and it is easy to misunderstand someone’s mental state, which can lead to negative consequences. Obtaining a deeper understanding of how social cognition changes across the lifespan is important because social interactions are an essential part of our lives; failure of these can lead to isolation and loneliness in all ages. Recent research has demonstrated that social development continues through adolescence and well into our twenties (e.g., Blakemore, 2008; Dumontheil et al., 2010) and that specific impairments in these abilities emerge with increasing age (e.g., Bailey & Henry, 2008; German & Hehman, 2006; Phillips et al., 2011). Older age is associated with reduced opportunity for social interaction and increased difficulties on social tasks.

This thesis aimed to investigate whether and how social cognitive abilities change across the lifespan, and whether we can enhance them by training the underlying cognitive functions. Across three empirical chapters I presented evidence from large samples of participants, comparing adolescents, young adults and older adults. An important strength of the work presented here is the broad range of tasks used to examine multiple sub-components of social cognition, including ecologically valid real-world measures of social attention and implicit measures of empathy for others' pain. The experiments included behavioural measures, as well as mobile eye-tracking and EEG, and the work in Chapters 2 and 3 was pre-registered to reduce the risk of publication and inference bias.

The conclusions can be summarised in three points: i) Age modulates different sub-components of social cognition in distinct ways, though more cognitively demanding components are more likely to suffer a decline in older age, ii) Cognitive training interventions can clearly bring direct improvements in the trained tasks, however near and far transfer of skills are limited, iii) Cognitive training has limited if any benefits to social cognition, even in those sub-components that are thought to rely most on cognitive abilities.

REFERENCES

Aboulafia-Brakha, T., Christe, B., Martory, M. D., & Annoni, J. M. (2011). Theory of mind tasks and executive functions: a systematic review of group studies in neurology. *Journal of Neuropsychology*, *5*(1), 39-55.

Adenzato, M., Brambilla, M., Manenti, R., De Lucia, L., Trojano, L., Garofalo, S., Enrici, I., & Cotelli, M. (2017). Gender differences in cognitive Theory of Mind revealed by transcranial direct current stimulation on medial prefrontal cortex. *Scientific Reports*, *7*(1), 1-9.

Adolph, R. (2009). The social brain: neural basis of social knowledge. *Annual review of psychology*, *60*, 693-716.

Allison, C., Auyeung, B., & Baron-Cohen, S. (2012). Toward brief "red flags" for autism screening: the short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls. *Journal of the American Academy of Child & Adolescent Psychiatry, 51*(2), 202-212.

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature reviews neuroscience*, *7*(4), 268-277.

Anderson, E., Clarke, L., & Spain, B. (2002). *Disability in adolescence*. Routledge.

Apperly, I. A. (2009). Alternative routes to perspective-taking: Imagination and rule-use may be better than simulation and theorising. *British Journal of Developmental Psychology*, *27*(3), 545.

Apperly, I. (2010). *Mindreaders: the cognitive basis of" theory of mind"*. Psychology Press.

Apperly, I. A., Back, E., Samson, D., & France, L. (2008). The cost of thinking about false beliefs: Evidence from adults' performance on a non-inferential theory of mind task. *Cognition*, *106*(3), 1093-1108.

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological review*, *116*(4), 953.

Apperly, I. A., Samson, D., Chiavarino, C., & Humphreys, G.W. (2004) Frontal and temporo-parietal lobe contributions to theory of mind: Neuropsychological evidence from a false-belief task with reduced language and executive demands. J*ournal of Cognitive Neuroscience, 16*, 1773–1784.

Apperly, I. A., Warren, F., Andrews, B. J., Grant, J., & Todd, S. (2011). Developmental continuity in theory of mind: Speed and accuracy of belief–desire reasoning in children and adults. *Child Development*, *82*, 1691–1703.

Appleton, M., & Reddy, V. (1996). Teaching three year-olds to pass false belief tests: A conversational approach. *Social Development*, *5*(3), 275-291.

Au, J., Sheehan, E., Tsai, N., Duncan, G. J., Buschkuehl, M., & Jaeggi, S. M. (2015). Improving fluid intelligence with training on working memory: a meta-analysis. *Psychonomic bulletin & review*, *22*(2), 366-377.

Baddeley, A. (1986). Modularity, mass-action and memory. *The Quarterly Journal of Experimental Psychology*, *38*(4), 527-533.

Bailey, P. E., Brady, B., Ebner, N. C., & Ruffman, T. (2020). Effects of age on emotion regulation, emotional empathy, and prosocial behavior. *The Journals of Gerontology: Series B*, *75*(4), 802-810.

Bailey, P. E., & Henry, J. D. (2008). Growing less empathic with age: Disinhibition of the self-perspective. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *63*(4), P219-P226.

Bailey, P. E., & Henry, J. D. (2010). Separating component processes of theory of mind in schizophrenia. *British Journal of Clinical Psychology, 49*(1), 43-52.

Bailey, P. E., Slessor, G., Rendell, P. G., Bennetts, R. J., Campbell, A., & Ruffman, T. (2014). Age differences in conscious versus subconscious social perception: The influence of face age and valence on gaze following. *Psychology and Aging*, *29*(3), 491.

Baker, L. J., Levin, D. T., & Saylor, M. M. (2016). The extent of default visual perspective taking in complex layouts. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(4), 508.

Ballesteros, S., Prieto, A., Mayas, J., Toril, P., Pita, C., Ponce de León, L., Reales, J.M., & Waterworth, J. (2014). Brain training with non-action video games enhances aspects of cognition in older adults: a randomized controlled trial. *Frontiers in aging neuroscience*, *6*, 277.

Ballesteros, S., Voelcker-Rehage, C., & Bherer, L. (2018). Cognitive and Brain Plasticity Induced by Physical Exercise, Cognitive Training, Video Games, and Combined Interventions. *Frontiers in human neuroscience, 12*, 169.

Banerjee, R., Watling, D., & Caputi, M. (2011). Peer relations and the understanding of faux pas: Longitudinal evidence for bidirectional associations. *Child development*, *82*(6), 1887-1905.

Barenberg, J., Berse, T., & Dutke, S. (2015). Ergometer cycling enhances executive control in task switching. *Journal of Cognitive Psychology*, *27*(6), 692-703.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn?: A taxonomy for far transfer. *Psychological bulletin*, *128*(4), 612.

Baron-Cohen S. (1997). Mindblindness. Cambridge, MA; London: The MIT Press

Barzy, M., Ferguson, H. J., & Williams, D. M. (2020). Perspective influences eye movements during real-life conversation: Mentalising about self vs. others in autism. *Autism*.

Bechi, M., Riccaboni, R., Ali, S., Fresi, F., Buonocore, M., Bosia, M., Cocchi, F., Smeraldi, E., & Cavallaro, R. (2012). Theory of mind and emotion processing training for patients with schizophrenia: preliminary findings. *Psychiatry research*, *198*(3), 371-377.

Beadle, J. N., & De la Vega, C. E. (2019). Impact of aging on empathy: Review of psychological and neural mechanisms. *Frontiers in psychiatry, 10,* 331.

Beelmann, A., & Heinemann, K. S. (2014). Preventing prejudice and improving intergroup attitudes: A meta-analysis of child and adolescent training programs. *Journal of applied developmental psychology*, *35*(1), 10-24.

Begeer, S., Gevers, C., Clifford, P., Verhoeve, M., Kat, K., Hoddenbach, E., & Boer, F. (2011). Theory of mind training in children with autism: A randomized controlled trial. *Journal of autism and developmental disorders*, *41*(8), 997-1006.

Béland, F., Zunzunegui, M. V., Alvarado, B., Otero, A., & Del Ser, T. (2005). Trajectories of cognitive decline and social relations. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 60*(6), P320-P330.

Bennett, I. J., Rivera, H. G., & Rypma, B. (2013). Isolating age-group differences in working memory load-related neural activity: Assessing the contribution of working memory capacity using a partial-trial fMRI method. *NeuroImage, 72,* 20–32.

Berkman, E. T., Kahn, L. E., & Merchant, J. S. (2014). Training-induced changes in inhibitory control network activity. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, *34*(1), 149–157.

Bernstein, D. M., Thornton, W. L., & Sommerville, J. A. (2011). Theory of mind through the ages: Older and middle-aged adults exhibit more errors than do younger adults on a continuous false belief task. *Experimental Aging Research*, *37*(5), 481-502.

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child development*, *81*(6), 1641-1660.

Bherer, L., Erickson, K. I., & Liu-Ambrose, T. (2013). A review of the effects of physical activity and exercise on cognitive and brain functions in older adults. *Journal of aging research*, *2013*.

Birmingham, E., Bischof, W. F., & Kingstone, A. (2009). Saliency does not account for fixations to eyes within social scenes. *Vision research*, *49*(24), 2992-3000.

Bishop, D., Aamodt-Leeper, G., Creswell, C., McGurk, R., & Skuse, D. (2001). Individual Differences in Cognitive Planning on the Tower of Hanoi Task: Neuropsychological Maturity or Measurement Error? *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42*(4), 551-556.

Blacker, K. J., Negoita, S., Ewen, J. B., & Courtney, S. M. (2017). N-back versus complex span working memory training. *Journal of Cognitive Enhancement*, *1*(4), 434-454.

Blakemore, S. J. (2008). The social brain in adolescence. *Nature Reviews Neuroscience*, *9*(4), 267-277.

Blakemore, S. J. (2010). The developing social brain: implications for education. *Neuron*, *65*(6), 744-747.

Blakemore, S. J., & Choudhury, S. (2006). Development of the adolescent brain: implications for executive function and social cognition. *Journal of child psychology and psychiatry*, *47*(3-4), 296-312.

Blakemore, S. J., den Ouden, H., Choudhury, S., & Frith, C. (2007). Adolescent development of the neural circuitry for thinking about intentions.

Blakemore, S. J., & Mills, K. L. (2014). Is adolescence a sensitive period for sociocultural processing?. *Annual review of psychology, 65*, 187-207.

Bock, A. M., Gallaway, K. C., & Hund, A. M. (2015). Specifying links between executive functioning and theory of mind during middle childhood: Cognitive flexibility predicts social understanding. *Journal of Cognition and Development*, *16*(3), 509-521.

Borella, E., Carretti, B., Riboldi, F., & De Beni, R. (2010). Working memory training in older adults: evidence of transfer and maintenance effects. *Psychology and aging, 25*(4), 767.

Bottiroli, S, Cavallini, E., Ceccato, I., Vecchi, T., & Lecce, S. (2016) "Theory of Mind in aging: Comparing cognitive and affective components in the faux pas test." *Archives of Gerontology and Geriatrics* 62: 152-162.

Bradford, E. E., Brunsdon, V. E., & Ferguson, H. J. (2020). The neural basis of belief-attribution across the lifespan: False-belief reasoning and the N400 effect. *Cortex, 126*, 265-280.

Bradford, E. E., Gomez, J. C., & Jentzsch, I. (2019). Exploring the role of self/other perspective-shifting in theory of mind with behavioural and EEG measures. *Social neuroscience, 14*(5), 530-544.

Bradford, E. E., Jentzsch, I., & Gomez, J. C. (2015). From self to social cognition: Theory of mind mechanisms and their relation to executive functioning. *Cognition*, *138*, 21-34.

Braver, T. S., Barch, D. M., & Cohen, J. D. (1999). Cognition and control in schizophrenia: a computational model of dopamine and prefrontal function. *Biological psychiatry*, *46*(3), 312-328.

Braver, T.S., & West, R. (2011). Working Memory, Executive Control and Aging. The Handbook of Aging and Cognition Third Edition.

Brehmer, Y., Shing, Y. L., Heekeren, H. R., Lindenberger, U., & Bäckman, L. (2016). Training-induced changes in subsequent-memory effects: No major differences among children, younger adults, and older adults. *NeuroImage*, *131*, 214-225.

Brehmer, Y., Westerberg, H., & Bäckman, L. (2012). Working-memory training in younger and older adults: training gains, transfer, and maintenance. *Frontiers in human neuroscience*, *6*, 63.

Brothers, L. (1990). The neural basis of primate social communication. *Motivation and emotion*, *14*(2), 81-91.

Brown-Schmidt, S. (2009). The role of executive function in perspective taking during online language comprehension. *Psychonomic bulletin & review*, *16*(5), 893-900.

Brown-Schmidt, S., & Fraundorf, S. H. (2015). Interpretation of informational questions modulated by joint knowledge and intonational contours. *Journal of Memory and Language*, *84*, 49-74.

Brüne, M. (2005a). “Theory of mind” in schizophrenia: a review of the literature. *Schizophrenia bulletin*, *31*(1), 21-42.

Brüne, M. (2005b). Emotion recognition,‘theory of mind,’and social behavior in schizophrenia. *Psychiatry research*, *133*(2-3), 135-147.

Brunsdon, V. E., Bradford, E. E., & Ferguson, H. J. (2019). Sensorimotor mu rhythm during action observation changes across the lifespan independently from social cognitive processes. *Developmental cognitive neuroscience*, *38*, 100659.

Brunsdon, V. E., Bradford, E. E., & Ferguson, H. J. (submitted). The developmental trajectories of executive function from adolescence to old age.

Buckner, R. L., & Louis, S. (2004). Memory and executive function review in aging and AD. *Neuron, 44*, 195–208.

Buitenweg, J. I., Murre, J. M., & Ridderinkhof, K. R. (2012). Brain training in progress: a review of trainability in healthy seniors. *Frontiers in human neuroscience*, *6*, 183.

Burnett, S., Bird, G., Moll, J., Frith, C., & Blakemore, S. J. (2009). Development during adolescence of the neural processing of social emotion. *Journal of cognitive neuroscience*, *21*(9), 1736-1750.

Burnett Heyes, S., Sebastian, C., Kadosh, K. C., & Blakemore, S. J. (2011). The social brain in adolescence: evidence from functional magnetic resonance imaging and behavioural studies. *Neuroscience & Biobehavioral Reviews*, *35*(8), 1654-1664.

Burnett Heyes, S., Zokaei, N., van der Staaij, I., Bays, P. M., & Husain, M. (2012). Development of visual working memory precision in childhood. *Developmental science*, *15*(4), 528-539.

Buschkuehl, M., Jaeggi, S. M., & Jonides, J. (2012). Neuronal effects following working memory training. *Developmental cognitive neuroscience*, *2*, S167-S179.

Buss, D. M., & Craik, K. H. (1983). The act frequency approach to personality. *Psychological Review*, *90*(2), 105.

Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition*, *112*(2), 337-342.

Butterfill, S. A., & Apperly, I. A. (2011). Monitoring and Controlling the Mental States of Others. *The Quarterly Journal of Experimental Psychology*, *63*, 1201-1217.

Cabeza, R. (2002). Hemispheric asymmetry reduction in older adults: the HAROLD model. *Psychology and aging*, *17*(1), 85.

Cane, J. E., Ferguson, H. J., & Apperly, I. A. (2017). Using perspective to resolve reference: The impact of cognitive load and motivation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(4), 591.

Cañigueral, R., & Hamilton, A. F. D. C. (2019). Being watched: Effects of an audience on eye gaze and prosocial behaviour. *Acta psychologica*, *195*, 50-63.

Cao, Y., Yusri, N. M., Powell, T., & Cunnington, R. (2019). Neural and behavioral markers of observed pain of older adults. *Neuropsychologia*, *131*, 84-90.

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental neuropsychology*, *28*(2), 595-616.

Carlson, S. M., Mandell, D. J., & Williams, L. (2004). Executive function and theory of mind: stability and prediction from ages 2 to 3. *Developmental Psychology*, *40*, 1105.

Carlson S.M. & Moses L.J. (2001) Individual differences in inhibitory control and children's theory of mind. *Child Development 72*(4): 1032–1053.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge University Press.

Carroll, J. B., & Maxwell, S. E. (1979). Individual differences in cognitive abilities. *Annual review of psychology*, *30*(1), 603-640.

Carruthers, P. (2016). Two systems for mindreading?. *Review of Philosophy and Psychology*, *7*(1), 141-162.

Carruthers, P. (2017). Mindreading in adults: evaluating two-systems views. *Synthese*, *194*(3), 673-688.

Carruthers, P., & Smith, P. K. (Eds.). (1996). *Theories of theories of mind*. Cambridge University Press.

Casey, B. J. (2015). Beyond simple models of self-control to circuit-based accounts of adolescent behavior. *Annual review of psychology*, *66*, 295-319.

Castelli, I., Baglio, F., Blasi, V., Alberoni, M., Falini, A., Liverta-Sempio, O., Nemni, R., & Marchetti, A. (2010). Effects of aging on mind reading ability through the eyes: an fMRI study. *Neuropsychologia 48*(9), 2586–2594.

Catmur, C., Walsh, V., & Heyes, C. (2007). Sensorimotor learning configures the human mirror system. *Current biology*, *17*(17), 1527-1531.

Cavallini, E., Bianco, F., Bottiroli, S., Rosi, A., Vecchi, T., & Lecce, S. (2015). Training for generalization in Theory of Mind: a study with older adults. *Frontiers in Psychology*, *6*, 1123.

Cavallini, E., Lecce, S., Bottiroli, S., Palladino, P., & Pagnin, A. (2013). Beyond false belief: Theory of mind in young, young-old, and old-old adults. *International Journal of Aging & Human Development*, *76*, 181–198.

Cepeda, N. J., Kramer, A. F., & Gonzalez de Sather, J. (2001). Changes in executive control across the life span: examination of task-switching performance. *Developmental psychology*, *37*(5), 715.

Charlton, R. A., Barrick, T. R., Markus, H. S., & Morris, R. G. (2009). Theory of mind associations with other cognitive functions and brain imaging in normal aging. *Psychology and aging*, *24*(2), 338.

Chatham, C. H., Claus, E. D., Kim, A., Curran, T., Banich, M. T., & Munakata, Y. (2012). Cognitive control reflects context monitoring, not motoric stopping, in response inhibition. *PloS one*, *7*(2), e31546.

Chein, J. M., & Morrison, A. B. (2010). Expanding the mind's workspace: Training and transfer effects with a complex working memory span task. *Psychonomic bulletin & review*, *17*(2), 193-199.

Chen, C., Yang, C. Y., & Cheng, Y. (2012). Sensorimotor resonance is an outcome but not a platform to anticipating harm to others. *Social Neuroscience*, *7*(6), 578-590.

Chen, Y. C., Chen, C. C., Decety, J., & Cheng, Y. (2014). Aging is associated with changes in the neural circuits underlying empathy. *Neurobiology of aging*, *35*(4), 827-836.

Cheng, Y., Chen, C., & Decety, J. (2014). An EEG/ERP investigation of the development of empathy in early and middle childhood. *Developmental Cognitive Neuroscience*, *10*, 160-169.

Cheng, Y., Lee, P. L., Yang, C. Y., Lin, C. P., Hung, D., & Decety, J. (2008). Gender differences in the mu rhythm of the human mirror-neuron system. *PLoS One*, *3*(5), e2113.

Cho, I., & Cohen, A. S. (2019). Explaining age-related decline in theory of mind: Evidence for intact competence but compromised executive function. *PloS one*, *14*(9), e0222890.

Choudhury, S., Blakemore, S. J., & Charman, T. (2006). Social cognitive development during adolescence. *Social cognitive and affective neuroscience*, *1*(3), 165-174.

Codispoti, M., Ferrari, V., & Bradley, M. M. (2007). Repetition and event-related potentials: distinguishing early and late processes in affective picture perception. *Journal of Cognitive Neuroscience*, *19*(4), 577-586.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd Ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, S., Bixenman, M., Meiran, N., & Diamond, A. (2001). *Task switching in children*. Presented at the South Carolina Bicentennial Symposium on Attention, University of South Carolina, Columbia, SC, USA.

Cohen, J. D., Forman, S. D., Braver, T. S., Casey, B. J., Servan-Schreiber, D., & Noll, D. C. (1993). Activation of the prefrontal cortex in a nonspatial working memory task with functional MRI. *Human Brain Mapping*, *1*, 293–304.

Colcombe, S. J., Erickson, K. I., Scalf, P. E., Kim, J. S., Prakash, R., McAuley, E., Elavsky, S., Marquez, D.X., Hu, L. & Kramer, A. F. (2006). Aerobic exercise training increases brain volume in aging humans. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *61*(11), 1166-1170.

Colcombe, S., & Kramer, A. F. (2003). Fitness effects on the cognitive function of older adults: a meta-analytic study. *Psychological science*, *14*(2), 125-130.

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user’s guide. *Psychonomic bulletin & review*, *12*(5), 769-786.

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron*, *58*(3), 306-324.

Cowan, N. (2011). The focus of attention as observed in visual working memory tasks: Making sense of competing claims. *Neuropsychologia*, *49*(6), 1401-1406.

Cowan, N., Morey, C. C., AuBuchon, A. M., Zwilling, C. E., & Gilchrist, A. L. (2010). Seven-year-olds allocate attention like adults unless working memory is overloaded. *Developmental science*, *13*(1), 120-133.

Craig, A. D. (2003). Interoception: the sense of the physiological condition of the body. *Current opinion in neurobiology*, *13*(4), 500-505.

Craig, A. D., & Craig, A. D. (2009). How do you feel--now? The anterior insula and human awareness. *Nature reviews neuroscience*, *10*(1).

Crone, E. A., Donohue, S. E., Honomichl, R., Wendelken, C., & Bunge, S. A. (2006). Brain regions mediating flexible rule use during development. *Journal of Neuroscience*, *26*(43), 11239-11247.

Crone, E. A., Wendelken, C., Donohue, S. E., & Bunge, S. A. (2005). Neural evidence for dissociable components of task-switching. *Cerebral cortex*, *16*(4), 475-486.

Dahlin, E., Nyberg, L., Bäckman, L., & Neely, A. S. (2008). Plasticity of executive functioning in young and older adults: immediate training gains, transfer, and long-term maintenance. *Psychology and aging*, *23*(4), 720.

Davidson, D. J., Zacks, R. T., & Williams, C. C. (2003). Stroop interference, practice, and aging. *Aging, Neuropsychology, and Cognition*, *10*(2), 85-98.

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, *44*(11), 2037–2078.

Davies, M., & Stone, T. (1995). *Mental Simulation: Evaluations and Applications-Reading in Mind and Language*. Oxford: Blackwell.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy.

De Jaegher, H., Di Paolo, E., & Gallagher, S. (2010). Can social interaction constitute social cognition?. *Trends in cognitive sciences*, *14*(10), 441-447.

De Luca, C. R., Wood, S. J., Anderson, V., Buchanan, J. A., Proffitt, T. M., Mahony, K., & Pantelis, C. (2003). Normative data from the CANTAB. I: development of executive function over the lifespan. *Journal of clinical and experimental neuropsychology*, *25*(2), 242-254.

Decety, J. (2010). The neurodevelopment of empathy in humans. *Developmental neuroscience*, *32*(4), 257-267.

Decety, J., & Michalska, K. J. (2010). Neurodevelopmental changes in the circuits underlying empathy and sympathy from childhood to adulthood. *Developmental science, 13*(6), 886-899.

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.

Deci, E. L., & Ryan, R. M. (2010). Intrinsic motivation. *The corsini encyclopedia of psychology*, 1-2.

Derwinger, A., Neely, A. S., Persson, M., Hill, R. D., & Bäckman, L. (2003). Remembering numbers in old age: Mnemonic training versus self-generated strategy training. *Aging, Neuropsychology, and Cognition*, *10*(3), 202-214.

Devine, R. T., & Hughes, C. (2013). Silent films and strange stories: Theory of mind, gender, and social experiences in middle childhood. *Child development*, *84*(3), 989-1003.

Diamond, A. (1985). Development of the ability to use recall to guide action, as indicated by infants' performance on AB. *Child development,* 868-883.

Diamond, A. (1995). Evidence of robust recognition memory early in life even when assessed by reaching behavior. *Journal of experimental child psychology*, *59*(3), 419-456.

Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. *Principles of frontal lobe function*, 466-503.

Diamond A. (2013). Executive functions. *Annual review of psychology, 64*, 135–168.

Diamond, A., & Ling, D. S. (2016). Conclusions about interventions, programs, and approaches for improving executive functions that appear justified and those that, despite much hype, do not. *Developmental cognitive neuroscience*, *18*, 34-48.

Doherty-Sneddon, G., Bruce, V., Bonner, L., Longbotham, S., & Doyle, C. (2002). Development of gaze aversion as disengagement from visual information. *Developmental psychology*, *38*(3), 438.

Doherty-Sneddon, G., & Phelps, F. G. (2005). Gaze aversion: A response to cognitive or social difficulty?. *Memory & cognition*, *33*(4), 727-733.

Döhnel, K., Schuwerk, T., Meinhardt, J., Sodian, B., Hajak, G., & Sommer, M. (2012). Functional activity of the right temporo-parietal junction and of the medial prefrontal cortex associated with true and false belief reasoning. *Neuroimage*, *60*(3), 1652-1661.

Dowiasch, S., Marx, S., Einhäuser, W., & Bremmer, F. (2015). Effects of aging on eye movements in the real world. *Frontiers in human neuroscience*, *9*, 46.

Dowsett, S. M., & Livesey, D. J. (2000). The development of inhibitory control in preschool children: Effects of “executive skills” training. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, *36*(2), 161-174.

Duval, C., Piolino, P., Bejanin, A., Eustache, F., & Desgranges, B. (2011). Age effects on different components of theory of mind. *Consciousness and Cognition*, *20*, 627–642.

Dumontheil, I., Apperly, I. A., & Blakemore, S. J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental science*, *13*(2), 331-338.

Eisenberger, N. I. (2011). Why rejection hurts: What social neuroscience has revealed about the brain’s response to social rejection. *Brain*, *3*(2), 1.

Eisenberg, N. I. (2012). The pain of social disconnection: Examining the shared neural underpinnings of social and physical pain. *Nature Reviews, Neuroscience, 13*, 421-34.

Eisenberger, N. I., & Lieberman, M. D. (2004). Why rejection hurts: a common neural alarm system for physical and social pain. *Trends in cognitive sciences*, *8*(7), 294-300.

Emery, N. J. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience & biobehavioral reviews, 24*(6), 581-604.

Enge, S., Behnke, A., Fleischhauer, M., Küttler, L., Kliegel, M., & Strobel, A. (2014). No evidence for true training and transfer effects after inhibitory control training in young healthy adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(4), 987.

Engle, R. W., & Kane, M. J. (2004). Executive Attention, Working Memory Capacity, and a Two-Factor Theory of Cognitive Control. In B. H. Ross (Ed.), The psychology of learning and motivation: Advances in research and theory, (Vol. 44, pp. 145–199). New York, NY: Elsevier Science

Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004a). Perspective taking as egocentric anchoring and adjustment. *Journal of personality and social psychology*, *87*(3), 327.

Epley, N., Morewedge, C. K., & Keysar, B. (2004b). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology*, *40*(6), 760-768.

Erickson, K. I., Voss, M. W., Prakash, R. S., Basak, C., Szabo, A., Chaddock, L., Kim, J.S., Heo, S.,Alves, H., White, S.M., & Wojcicki, T. R. (2011). Exercise training increases size of hippocampus and improves memory. *Proceedings of the National Academy of Sciences, 108*(7), 3017-3022.

Espy, K. A. (1997). The Shape School: Assessing executive function in preschool children. *Developmental Neuropsychology*, *13*(4), 495-499.

Ferguson, H. J., & Breheny, R. (2011). Eye movements reveal the time-course of anticipating behaviour based on complex, conflicting desires. *Cognition*, *119*(2), 179-196.

Ferguson, H. J., & Breheny, R. (2012). Listeners' eyes reveal spontaneous sensitivity to others' perspectives. *Journal of Experimental Social Psychology*, *48*(1), 257-263.

Ferguson, H. J., Cane, J. E., Douchkov, M., & Wright, D. (2015). Empathy predicts false belief reasoning ability: evidence from the N400. *Social Cognitive and Affective Neuroscience*, *10*(6), 848-855.

Ferguson, H. J., & Cane, J. (2017). Tracking the impact of depression in a perspective-taking task. *Scientific Reports*, *7*(1), 1-9.

Ferguson, H., Brunsdon, V., & Bradford, E. (2018). Age of avatar modulates the altercentric bias in a visual perspective-taking task: ERP and behavioural evidence. *Cognitive, Affective, and Behavioral Neuroscience, 18*, 1298-1319.

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of experimental psychology: General*, *137*(2), 201.

Fisher, N., & Happé, F. (2005). A training study of theory of mind and executive function in children with autistic spectrum disorders. *Journal of autism and developmental disorders*, *35*(6), 757.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition*. Mcgraw-Hill Book Company.

Fiske S.T., & Taylor S.E. (2013). *Social Cognition: From Brains to Culture*. London: Sage.

Flavell, J. H. (2004). Theory-of-mind development: Retrospect and prospect. *Merrill-Palmer Quarterly (1982-),* 274-290.

Flavell, J. H., Everett, B. a., Croft, K., & Flavell, E. R. (1981). Young children's knowledge about visual perception: Further evidence for the Level 1-Level 2 distinction. *Developmental Psychology, 17*(1) 99-103.

Fosco, W. D., Hawk Jr, L. W., Colder, C. R., Meisel, S. N., & Lengua, L. J. (2019). The development of inhibitory control in adolescence and prospective relations with delinquency. *Journal of adolescence, 76,* 37-47.

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.

Foulsham, T., & Kingstone, A. (2017). Are fixations in static natural scenes a useful predictor of attention in the real world?. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 71*(2), 172.

Foulsham, T., Walker, E., & Kingstone, A. (2011). The where, what and when of gaze allocation in the lab and the natural environment. *Vision research*, *51*(17), 1920-1931.

Foroughi, C. K., Monfort, S. S., Paczynski, M., McKnight, P. E., & Blaood, P. M. (2016). Placebo effects in cognitive training. *Proceedings of the national Academy of Sciences*, *113*(27), 7470-7474.

Fraser, A. M., Hampton, R. S., Spinrad, T. L., Varnum, M., Blais, C., Eisenberg, N., Gal-Szabo, D.E., Berger, R.H., Xu, J. & Xiao, S.X. (2020). Children’s mu suppression is sensitive to witnessing others’ social victimization. *Social Neuroscience*, 1-7.

Freeth, M., Foulsham, T., & Kingstone, A. (2013). What affects social attention? Social presence, eye contact and autistic traits. *PloS one*, *8*(1), e53286.

Frick, A., Möhring, W., & Newcombe, N. S. (2014). Picturing perspectives: development of perspective-taking abilities in 4-to 8-year-olds. *Frontiers in psychology*, *5*, 386.

Frith C. D. (2008). Social cognition. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *363*(1499), 2033–2039.

Frith, C. D., & Frith, U. (2007). Social cognition in humans. *Current biology*, *17*(16), R724-R732.

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1431), 459-473.

Furlanetto, T., Becchio, C., Samson, D., & Apperly, I. (2016). Altercentric interference in level 1 visual perspective taking reflects the ascription of mental states, not submentalizing. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(2), 158.

Gallagher, S., & Hutto, D. (2008). Understanding others through primary interaction and narrative practice. *The shared mind: Perspectives on intersubjectivity*, *12*, 17-38.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*(2), 593-609.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences, 2*(12), 493–501.

Gamboz, N., Borella, E., & Brandimonte, M. A. (2009). The role of switching, inhibition and working memory in older adults' performance in the Wisconsin Card Sorting Test. *Aging, Neuropsychology, and Cognition*, *16*(3), 260-284.

Gao, Z., Ye, T., Shen, M., & Perry, A. (2016). Working memory capacity of biological movements predicts empathy traits. *Psychonomic Bulletin & Review*, *23*(2), 468-475.

Gathercole, S. E., Dunning, D. L., Holmes, J., & Norris, D. (2019). Working memory training involves learning new skills. *Journal of Memory and Language*, *105*, 19-42.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental psychology*, *40*(2), 177.

Gazzaley, A., Cooney, J. W., McEvoy, K., Knight, R. T., & D'esposito, M. (2005). Top-down enhancement and suppression of the magnitude and speed of neural activity. *Journal of cognitive neuroscience*, *17*(3), 507-517.

Gbadeyan, O., McMahon, K., Steinhauser, M., & Meinzer, M. (2016). Stimulation of dorsolateral prefrontal cortex enhances adaptive cognitive control: a high-definition transcranial direct current stimulation study. *Journal of Neuroscience*, *36*(50), 12530-12536.

Gbadeyan, O., Steinhauser, M., Hunold, A., Martin, A. K., Haueisen, J., & Meinzer, M. (2019). Modulation of adaptive cognitive control by prefrontal high-definition transcranial direct current stimulation in older adults. *The Journals of Gerontology: Series B*, *74*(7), 1174-1183.

Ge, Y., Grossman, R. I., Babb, J. S., Rabin, M. L., Mannon, L. J., & Kolson, D. L. (2002). Age-related total gray matter and white matter changes in normal adult brain. Part I: volumetric MR imaging analysis. *American journal of neuroradiology*, *23*(8), 1327-1333.

German, T. P., & Hehman, J. A. (2006). Representational and executive selection resources in ‘theory of mind’: Evidence from compromised belief-desire reasoning in old age. *Cognition*, *101*(1), 129-152.

Gheysen, F., Poppe, L., DeSmet, A., Swinnen, S., Cardon, G., De Bourdeaudhuij, I., Chastin, S., & Fias, W. (2018). Physical activity to improve cognition in older adults: can physical activity programs enriched with cognitive challenges enhance the effects? A systematic review and meta-analysis. *International Journal of Behavioral Nutrition and Physical Activity*, *15*(1), 63.

Glenberg, A. M., Schroeder, J. L., & Robertson, D. A. (1998). Averting the gaze disengages the environment and facilitates remembering. *Memory & cognition*, *26*(4), 651-658.

Golan, O., & Baron-Cohen, S. (2006). Systemizing empathy: Teaching adults with Asperger syndrome or high-functioning autism to recognize complex emotions using interactive multimedia. *Development and psychopathology*, *18*(2), 591-617.

Goldstein, T. R., & Winner, E. (2010). Engagement in role play, pretense, and acting classes predict advanced theory of mind skill in middle childhood. *Imagination, Cognition and Personality*, *30*(3), 249-258.

Goldstein, T. R., Wu, K., & Winner, E. (2009). Actors are skilled in theory of mind but not empathy. *Imagination, Cognition and Personality*, *29*(2), 115-133.

Grainger, S. A., Steinvik, H. R., Henry, J. D., & Phillips, L. H. (2019). The role of social attention in older adults' ability to interpret naturalistic social scenes. *Quarterly Journal of Experimental Psychology*, *72*(6), 1328-1343.

Grant, D. A., & Berg, E. (1948). A behavioural analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of experimental psychology*, *38*(4), 404.

Green, M. F., Horan, W. P., & Lee, J. (2015). Social cognition in schizophrenia. *Nature Reviews Neuroscience*, *16*(10), 620-631.

Greenwood, P. M. (2007). Functional plasticity in cognitive aging: review and hypothesis. *Neuropsychology*, *21*(6), 657.

Guay, S., Héroux, J., Joubert, S., Béatrice, P., Koninck, D., Blanchette, V. L., & De Beaumont, L. (2018). Effects of age and sex on neurophysiological substrates of empathy: a cross-sectional EEG study.

Gutsell, J. N., & Inzlicht, M. (2010). Empathy constrained: Prejudice predicts reduced mental simulation of actions during observation of outgroups. *Journal of experimental social psychology*, *46*(5), 841-845.

Hallock, H., Collins, D., Lampit, A., Deol, K., Fleming, J., & Valenzuela, M. (2016). Cognitive training for post-acute traumatic brain injury: a systematic review and meta-analysis. *Frontiers in human neuroscience*, *10*, 537.

Han, K., Chapman, S. B., & Krawczyk, D. C. (2018). Neuroplasticity of cognitive control networks following cognitive training for chronic traumatic brain injury. *NeuroImage: Clinical*, *18*, 262-278.

Happé, F., Cook, J. L., & Bird, G. (2017). The structure of social cognition: In (ter) dependence of sociocognitive processes. *Annual Review of Psychology*, *68*, 243-267.

Happé, F., & Frith, U. (1996). Theory of mind and social impairment in children with conduct disorder. *British Journal of Developmental Psychology*, *14*(4), 385-398.

Happé, F., & Frith, U. (2014). Annual research review: Towards a developmental neuroscience of atypical social cognition. *Journal of Child Psychology and Psychiatry*, *55*(6), 553-577.

Harrison, T. L., Shipstead, Z., Hicks, K. L., Hambrick, D. Z., Redick, T. S., & Engle, R. W. (2013). Working memory training may increase working memory capacity but not fluid intelligence. *Psychological Science*, *24*(12), 2409-2419.

Hartman, M., Bolton, E., & Fehnel, S. E. (2001). Accounting for age differences on the Wisconsin Card Sorting Test: Decreased working memory, not inflexibility. *Psychology and aging*, *16*(3), 385.

Hartshorne, J. K., & Germine, L. T. (2015). When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the life span. *Psychological science*, *26*(4), 433-443.

Hartwright, C. E., Apperly, I. A., & Hansen, P. C. (2012). Multiple roles for executive control in belief–desire reasoning: Distinct neural networks are recruited for self perspective inhibition and complexity of reasoning. *NeuroImage*, *61*(4), 921-930.

Hasher, L., Lustig, C., & Zacks, R. T. (2007). Inhibitory mechanisms and the control of attention. In A. A. Conway, C. Jarrold, M. J. Kane, A. Miyake & J. N. Towse (Eds.), *Variation in working memory* (pp.227–249). Oxford, U.K.: Oxford University Press.

Hayward, D. A., Voorhies, W., Morris, J. L., Capozzi, F., & Ristic, J. (2017). Staring reality in the face: A comparison of social attention across laboratory and real-world measures suggests little common ground. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *71*(3), 212.

Hein, G., Silani, G., Preuschoff, K., Batson, C. D., & Singer, T. (2010). Neural responses to ingroup and outgroup members' suffering predict individual differences in costly helping. *Neuron*, *68*(1), 149-160.

Heinzel, S., Lorenz, R. C., Pelz, P., Heinz, A., Walter, H., Kathmann, N., Rapp, M.A., & Stelzel, C. (2016). Neural correlates of training and transfer effects in working memory in older adults. *Neuroimage*, *134*, 236-249.

Heinzel, S., Schulte, S., Onken, J., Duong, Q. L., Riemer, T. G., Heinz, A., Kathmann, N., & Rapp, M. A. (2014). Working memory training improvements and gains in non-trained cognitive tasks in young and older adults. *Aging, Neuropsychology, and Cognition*, *21*(2), 146-173.

Henry, J. D., Phillips, L. H., Ruffman, T., & Bailey, P. E. (2013). A meta-analytic review of age differences in theory of mind. *Psychology and Aging*, *28*(3), 826.

Henry, J. D., Phillips, L. H., & Von Hippel, C. (2014). A meta-analytic review of theory of mind difficulties in behavioural-variant frontotemporal dementia. *Neuropsychologia*, *56*, 53-62.

Hess, T. M., Osowski, N. L., & Leclerc, C. M. (2005). Age and experience influences on the complexity of social inferences. *Psychology and aging, 20*(3), 447.

Hobson, H. M., & Bishop, D. V. (2017). The interpretation of mu suppression as an index of mirror neuron activity: past, present and future. *Royal Society Open Science, 4*(3), 160662.

Holmes, J., Woolgar, F., Hampshire, A., & Gathercole, S. E. (2019). Are working memory training effects paradigm-specific? *Frontiers in psychology, 10*, 1103.

Hooker, C. I., Bruce, L., Fisher, M., Verosky, S. C., Miyakawa, A., & Vinogradov, S. (2012). Neural activity during emotion recognition after combined cognitive plus social cognitive training in schizophrenia. *Schizophrenia research, 139*(1-3), 53-59.

Hughes, C. (1998). Finding your marbles: Does preschoolers' strategic behavior predict later understanding of mind? *Developmental Psychology, 34*(6), 1326.

Hughes, C. (2011). Changes and challenges in 20 years of research into the development of executive functions. *Infant and Child Development, 20*(3), 251-271.

Hughes, C., & Ensor, R. (2007). Executive function and theory of mind: Predictive relations from ages 2 to 4. *Developmental Psychology, 43*, 1447–1459

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia, 44*(11), 2017.

Isaacowitz, D. M., & Stanley, J. T. (2011). Bringing an ecological perspective to the study of aging and recognition of emotional facial expressions: Past, current, and future methods. *Journal of nonverbal behavior, 35*(4), 261.

Jackson, P. L., Meltzoff, A. N., & Decety, J. (2005). How do we perceive the pain of others? A window into the neural processes involved in empathy. *Neuroimage*, *24*(3), 771-779.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences*, *105*(19), 6829-6833.

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the N-back task as a working memory measure. *Memory*, *18*(4), 394-412.

Johnstone, S. J., Dimoska, A., Smith, J. L., Barry, R. J., Pleffer, C. B., Chiswick, D., & Clarke, A. R. (2007). The development of stop-signal and Go/Nogo response inhibition in children aged 7–12 years: performance and event-related potential indices. *International Journal of Psychophysiology*, *63*(1), 25-38.

Jolles, D. D., van Buchem, M. A., Crone, E. A., & Rombouts, S. A. (2013). Functional brain connectivity at rest changes after working memory training. *Human brain mapping*, *34*(2), 396-406.

Kalish, C. W., & Shiverick, S. M. (2004). Children’s reasoning about norms and traits as motives for behavior. *Cognitive Development, 19*, 401–416.

Kanske, P., & Murray, R. J. (2019). Understanding others: The neurobiology of social cognition. *Cortex; a journal devoted to the study of the nervous system and behavior*, *121*, A1.

Karbach, J., & Kray, J. (2009). How useful is executive control training? Age differences in near and far transfer of task-switching training. *Developmental science, 12*(6), 978-990.

Karbach, J., & Verhaeghen, P. (2014). Making working memory work: a meta-analysis of executive-control and working memory training in older adults. *Psychological science*, *25*(11), 2027-2037.

Keenan, T., Olson, D. R., & Marini, Z. (1998). Working memory and children's developing understanding of mind. *Australian Journal of Psychology*, *50*(2), 76-82.

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior research methods*, *42*(3), 627-633.

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.

Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition*, *89*(1), 25-41.

Kingstone, A. (2009). Taking a real look at social attention. *Current opinion in neurobiology*, *19*(1), 52-56.

Kirkham, N. Z., Cruess, L., & Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimension-switching task. *Developmental science*, *6*(5), 449-467.

Klencklen, G., Lavenex, P. B., Brandner, C., & Lavenex, P. (2017). Working memory decline in normal aging: Memory load and representational demands affect performance. *Learning and Motivation*, *60*, 10-22.

Klimecki, O., & Singer, T. (2012). Empathic distress fatigue rather than compassion fatigue? Integrating findings from empathy research in psychology and social neuroscience. *Pathological altruism*, 368-383.

Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Training of working memory in children with ADHD. *Journal of clinical and experimental neuropsychology*, *24*(6), 781-791.

Klindt, D., Devaine, M., & Daunizeau, J. (2017). Does the way we read others' mind change over the lifespan? Insights from a massive web poll of cognitive skills from childhood to late adulthood. *Cortex*, *86*, 205-215.

Kloo, D., & Perner, J. (2003). Training transfer between card sorting and false belief understanding: Helping children apply conflicting descriptions. *Child development*, *74*(6), 1823-1839.

Kloo, D., & Perner, J. (2005). Disentangling dimensions in the dimensional change card-sorting task. *Developmental Science*, *8*(1), 44-56.

Kovács, Á. M., Téglás, E., & Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science*, *330*(6012), 1830-1834.

Krach, S., Cohrs, J. C., de Echeverría Loebell, N. C., Kircher, T., Sommer, J., Jansen, A., & Paulus, F. M. (2011). Your flaws are my pain: Linking empathy to vicarious embarrassment. *PLoS One*, *6*(4), e18675.

Krach, S., Kamp-Becker, I., Einhäuser, W., Sommer, J., Frässle, S., Jansen, A., Rademacher, F., S., Einhäuser, W., Müller-Pinzler, L., Gazzola, V., & Paulus, F. M. (2015). Evidence from pupillometry and fMRI indicates reduced neural response during vicarious social pain but not physical pain in autism. *Human Brain Mapping*, *36*(11), 4730-4744.

Kraft, E. (2012). Cognitive function, physical activity, and aging: possible biological links and implications for multimodal interventions. *Aging, Neuropsychology, and Cognition*, *19*(1-2), 248-263.

Kray, J., Eber, J., & Lindenberger, U. (2004). Age differences in executive functioning across the lifespan: The role of verbalization in task preparation. *Acta Psychologica*, *115*(2-3), 143-165.

Kross, E., Berman, M. G., Mischel, W., Smith, E. E., & Wager, T. D. (2011). Social rejection shares somatosensory representations with physical pain. *Proceedings of the National Academy of Sciences*, *108*(15), 6270-6275.

Lamm, C., Decety, J., & Singer, T. (2011). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *Neuroimage*, *54*(3), 2492-2502.

Lamm, C., Meltzoff, A. N., & Decety, J. (2010). How do we empathize with someone who is not like us? A functional magnetic resonance imaging study. *Journal of cognitive neuroscience*, *22*(2), 362-376.

Lang, B., & Perner, J. (2002). Understanding of intention and false belief and the development of self-control. *British Journal of Developmental Psychology*, *20,* 67–76.

Lecce, S., Bianco, F., Devine, R. T., Hughes, C., & Banerjee, R. (2014). Promoting theory of mind during middle childhood: A training program. *Journal of Experimental Child Psychology*, *126*, 52-67.

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.

Lecce, S., Ceccato, I., & Cavallini, E. (2019). Investigating ToM in aging with the MASC: from accuracy to error type. *Aging, Neuropsychology, and Cognition*, *26*(4), 541-557.

Lecce, S., Ceccato, I., Rosi, A., Bianco, F., Bottiroli, S., & Cavallini, E. (2019). Theory of mind plasticity in aging: The role of baseline, verbal knowledge, and executive functions. *Neuropsychological rehabilitation*, *29*(3), 440-455.

Leslie, A. M. (2005). Developmental parallels in understanding minds and bodies. *Trends in Cognitive Sciences*, *9*(10), 459–462.

Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in "theory of mind". *Trends in Cognitive Sciences*, *8*(12), 528–533.

Leśniak, M. M., Iwański, S., Szutkowska-Hoser, J., & Seniów, J. (2019). Comprehensive cognitive training improves attention and memory in patients with severe or moderate traumatic brain injury. *Applied Neuropsychology: Adult*, 1-10.

Levine, B., Stuss, D.T., Winocur, G., Binns, M.A., Fahy, L., Mandic, M., Bridges, K., & Robertson, I.H. (2007). Cognitive rehabilitation in the elderly: effects on strategic behavior in relation to goal management. *Journal of the International neuropsychological Society*, 13: 143-152.

Levy, J., & Feldman, R. (2017). Can teenagers feel the pain of others? Peeking into the teenage brain to find empathy. *Front Young Minds*, *5*, 1-8.

Li, K. Z., Roudaia, E., Lussier, M., Bherer, L., Leroux, A., & McKinley, P. A. (2010). Benefits of cognitive dual-task training on balance performance in healthy older adults. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, *65*(12), 1344-1352.

Lin, S., Keysar, B., & Epley, N. (2010). Reflexively mindblind: Using theory of mind to interpret behavior requires effortful attention. *Journal of Experimental Social Psychology*, *46*(3), 551-556.

Logan, G. D. (1994). On the ability to inhibit thought and action: A users' guide to the stop signal paradigm.

Long, M. R., Horton, W. S., Rohde, H., & Sorace, A. (2018). Individual differences in switching and inhibition predict perspective-taking across the lifespan. *Cognition*, *170*, 25-30.

Longman, C. S., Lavric, A., & Monsell, S. (2016). The coupling between spatial attention and other components of task-set: A task-switching investigation. *The Quarterly Journal of Experimental Psychology*, *69*(11), 2248-2275.

Loosli, S. V., Buschkuehl, M., Perrig, W. J., & Jaeggi, S. M. (2012). Working memory training improves reading processes in typically developing children. *Child Neuropsychology, 18*(1), 62-78.

Lövdén, M., Bäckman, L., Lindenberger, U., Schaefer, S., & Schmiedek, F. (2010). A theoretical framework for the study of adult cognitive plasticity. *Psychological bulletin*, *136*(4), 659.

Lövdén, M., Brehmer, Y., Li, S. C., & Lindenberger, U. (2012). Training induced compensation versus magnification of individual differences in memory performance. *Frontiers in Human Neuroscience, 6*, 1–14.

Luna, B., Thulborn, K.R., Munoz, D.P., Merriam, E.P., Garver, K.E., Minshew, N.J., Keshavan, M.S., Genovese, C.R., Eddy, W.F., & Sweeney, J.A., (2001). Maturation of widely distributed brain function subserves cognitive development. *Neuroimage, 13*(5), pp.786-793.

Lustig, C., Shah, P., Seidler, R., & Reuter-Lorenz, P. A. (2009). Aging, training, and the brain: a review and future directions. *Neuropsychology review*, *19*(4), 504-522.

Madden, D.J., Costello, M.C., Dennis, N.A., Davis, S.W., Shepler, A.M., Spaniol, J., Bu cur, B., & Cabeza, R. (2010). Adult age differences in functional connectivity during executive control. *Neuroimage*, *52*(2), 643-657.

Magar, E. C., Phillips, L. H., & Hosie, J. A. (2010). Brief report: Cognitive-regulation across the adolescent years. *Journal of Adolescence*, *33*(5), 779-781.

Mahy, C. E., Moses, L. J., & Pfeifer, J. H. (2014). How and where: Theory-of-mind in the brain. *Developmental cognitive neuroscience*, *9*, 68-81.

Mahy, C. E., Vetter, N., Kuhn-Popp, N., Locher, C., Krautschuk, S., & Kliegel, M. (2014b). The influence of inhibitory processes on affective theory of mind in young and old adults. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, 21, 129–145.

Malle, B. F., & Holbrook, J. (2012). Is there a hierarchy of social inferences? The likelihood and speed of inferring intentionality, mind, and personality. *Journal of Personality and Social Psychology*, *102*(4), 661.

Marangoni, C., Garcia, S., Ickes, W., & Teng, G. (1995). Empathic accuracy in a clinically relevant setting. *Journal of personality and social psychology, 68*(5), 854.

Maraver, M. J., Bajo, M. T., & Gomez-Ariza, C. J. (2016). Training on working memory and inhibitory control in young adults. *Frontiers in human neuroscience*, *10,* 588.

Martin, A. K., Huang, J., Hunold, A., & Meinzer, M. (2017). Sex mediates the effects of high-definition transcranial direct current stimulation on "mind-reading". *Neuroscience*, *366*, 84-94.

Martin, A., Huang, J., Matuschka, L., & Meinzer, M. (2019). Perspective switching in young and older adults.

Masangkay, Z. S., McCluskey, K. A., McIntyre, C. W., Sims-Knight, J., Vaughn, B. E., & Flavell, J. H. (1974). The early development of inferences about the visual percepts of others. *Child development*, 357-366.

Mattan, B. D., Quinn, K. A., Acaster, S. L., Jennings, R. M., & Rotshtein, P. (2017). Prioritization of self-relevant perspectives in ageing. *Quarterly Journal of Experimental Psychology, 70*(6), 1033-1052.

Maylor, E. A., Birak, K. S., & Schlaghecken, F. (2011). Inhibitory motor control in old age: evidence for de-automatization? *Frontiers in Psychology*, *2*, 132.

Maylor, E. A., Moulson, J. M., Muncer, A. M., & Taylor, L. A. (2002). Does performance on theory of mind tasks decline in old age? *British Journal of Psychology*, *93*(4), 465-485.

McKendrick, R., Ayaz, H., Olmstead, R., & Parasuraman, R. (2014). Enhancing dual-task performance with verbal and spatial working memory training: continuous monitoring of cerebral hemodynamics with NIRS. *Neuroimage*, *85*, 1014-1026.

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental psychology*, *49*(2), 270.

Melby-Lervåg, M., Redick, T. S., & Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of “far transfer” evidence from a meta-analytic review. *Perspectives on Psychological Science, 11*(4), 512-534.

Melinder, A., Gredebäck, G., Westerlund, A., & Nelson, C. A. (2010). Brain activation during upright and inverted encoding of own-and other-age faces: ERP evidence for an own-age bias. *Developmental science*, *13*(4), 588-598.

Mella, N., Studer, J., Gilet, A. L., & Labouvie-Vief, G. (2012). Empathy for pain from adolescence through adulthood: an event-related brain potential study. *Frontiers in Psychology*, *3*, 501.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: evidence of a dependence between retrieval operations. *Journal of experimental psychology*, *90*(2), 227.

Meyer, M. L., Taylor, S. E., & Lieberman, M. D. (2015). Social working memory and its distinctive link to social cognitive ability: an fMRI study. *Social cognitive and affective neuroscience*, *10*(10), 1338-1347.

Mills, K. L., Dumontheil, I., Speekenbrink, M., & Blakemore, S. J. (2015). Multitasking during social interactions in adolescence and early adulthood. *Royal Society Open Science*, *2*(11), 150117.

Minear, M., Brasher, F., Guerrero, C. B., Brasher, M., Moore, A., & Sukeena, J. (2016). A simultaneous examination of two forms of working memory training: Evidence for near transfer only. *Memory & Cognition*, *44*(7), 1014-1037.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive psychology*, *41*(1), 49-100.

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H, Houts, R., Poulton, R., Roberts, B.W., Ross, S., & Sears, M. R. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the national Academy of Sciences*, *108*(7), 2693-2698.

Moll, H., & Tomasello, M. (2004). 12-and 18-month-old infants follow gaze to spaces behind barriers. *Developmental science*, *7*(1), F1-F9.

Moskowitz, G. B., & Okten, I. O. (2017). Social Cognition. In *Getting Grounded in Social Psychology* (pp. 37-78). Psychology Press.

Moran, J. M., Jolly, E., & Mitchell, J. P. (2012). Social-cognitive deficits in normal aging. *Journal of neuroscience*, *32*(16), 5553-5561.

Narayanan, N. S., Prabhakaran, V., Bunge, S. A., Christoff, K., Fine, E. M., & Gabrieli, J. D. (2005). The role of the prefrontal cortex in the maintenance of verbal working memory: an event-related FMRI analysis. *Neuropsychology*, *19*(2), 223.

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.

Nettle, D. (2006). The evolution of personality variation in humans and other animals. *American Psychologist*, *61*(6), 622.

Niendam, T. A., Laird, A. R., Ray, K. L., Dean, Y. M., Glahn, D. C., & Carter, C. S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cognitive, Affective, & Behavioral Neuroscience*, *12*(2), 241-268.

Nilsson, L. G., Sternäng, O., Rönnlund, M., & Nyberg, L. (2009). Challenging the notion of an early-onset of cognitive decline. *Neurobiology of aging*, *30*(4), 521-524.

Noack, H., Lövdén, M., Schmiedek, F., & Lindenberger, U. (2009). Cognitive plasticity in adulthood and old age: Gauging the generality of cognitive intervention effects. *Restorative neurology and neuroscience*, *27*(5), 435-453.

Nouchi, R., Taki, Y., Takeuchi, H., Hashizume, H., Akitsuki, Y., Shigemune, Y, Sekiguchi, A., Kotozaki, Y., Tsukiura, T., Yomogida, Y., & Kawashima, R. (2012). Brain training game improves executive functions and processing speed in the elderly: a randomized controlled trial. *PloS one*, *7*(1), e29676.

O'Hare, E. D., & Sowell, E. R. (2008). Imaging developmental changes in gray and white matter in the human brain. *Handbook of developmental cognitive neuroscience*, *2*, 23-38.

Oberauer, K., Süß, H. M., Schulze, R., Wilhelm, O., & Wittmann, W. W. (2000). Working memory capacity facets of a cognitive ability construct. *Personality and individual differences*, *29*(6), 1017-1045.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs?. *Science*, *308*(5719), 255-258.

Orellana, G., & Slachevsky, A. (2013). Executive functioning in schizophrenia. *Frontiers in psychiatry*, *4*, 35.

Ouellet, J., Scherzer, P.B., Rouleau, I., Métras, P., Bertrand-Gauvin, C., Djerroud, N., Boisseau, E., & Duquette, P. (2010). Assessment of social cognition in patients with multiple sclerosis. *Journal of the International Neuropsychological Society*, *16*, 287-296.

Owen, A. M., Hampshire, A., Grahn, J. A., Stenton, R., Dajani, S., Burns, A. S., & Ballard, C. G. (2010). Putting brain training to the test. *Nature, 465*, 775–778.

Pacherie, E. (1997). *Motor-images, self-consciousness, and autism*. In J. Russell (Ed.), Autism as an executive disorder (p. 215-255). Oxford University Press.

Pardini, M., & Nichelli, P.F. (2009). Age-related decline in mentalizing skills across adult life span. *Experimental Aging Research, 35,* 98–106.

Park, D. C., & Bischof, G. N. (2013). The aging mind: neuroplasticity in response to cognitive training. *Dialogues in clinical neuroscience*, *15*(1), 109.

Pellicano, E. (2007). Links between theory of mind and executive function in young children with autism: Clues to developmental primacy. *Developmental psychology*, *43*(4), 974.

Peper, J. S., & Dahl, R. E. (2013). The teenage brain: Surging hormones—Brain-behavior interactions during puberty. *Current directions in psychological science*, *22*(2), 134-139.

Perner, J. (1998). The meta-intentional nature of executive functions and theory of mind. In P. Carruthers, & J. Boucher (Eds.), Language and thought: Interdisciplinary themes (pp. 270–283). Cambridge, England: Cambridge University Press.

Perner, J., & Lang, B. (1999). Development of theory of mind and executive control. *Trends in Cognitive Sciences, 3,* 337–344.

Perner, J., & Lang, B. (2000). Theory of mind and executive function: Is there a developmental relationship?

Perner, J., & Lang, B. (2002). What causes 3-year-olds' difficulty on the dimensional change card sorting task?. *Infant and Child Development: An International Journal of Research and Practice*, *11*(2), 93-105.

Perry, A., Bentin, S., Bartal, I. B. A., Lamm, C., & Decety, J. (2010). “Feeling” the pain of those who are different from us: Modulation of EEG in the mu/alpha range. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(4), 493-504.

Peters, R. (2006). Ageing and the brain. *Postgraduate medical journal*, *82*(964), 84-88.

Phillips, L. H., Bull, R., Allen, R., Insch, P., Burr, K., & Ogg, W. (2011). Lifespan aging and belief reasoning: Influences of executive function and social cue decoding. *Cognition*, *120*(2), 236-247.

Phillips, L. H., MacLean, R. D., & Allen, R. (2002). Age and the understanding of emotions: Neuropsychological and sociocognitive perspectives. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *57*(6), P526-P530.

Pineda, J. A. (2005). The functional significance of mu rhythms: translating "seeing" and "hearing" into "doing". *Brain research reviews*, *50*(1), 57-68.

Poletti, M., Enrici, I., & Adenzato, M. (2012). Cognitive and affective theory of mind in neurodegenerative diseases: neuropsychological, neuroanatomical and neurochemical levels. *Neuroscience & Biobehavioral Reviews, 36*(9), 2147–2164.

Preis, M. A., Kröner-Herwig, B., Schmidt-Samoa, C., Dechent, P., & Barke, A. (2015). Neural correlates of empathy with pain show habituation effects. An fMRI study. *PloS one*, *10*(8), e0137056.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioural and Brain Sciences*, *1*(4), 515–526.

Prencipe, A., Kesek, A., Cohen, J., Lamm, C., Lewis, M. D., & Zelazo, P. D. (2011). Development of hot and cool executive function during the transition to adolescence. *Journal of experimental child psychology*, *108*(3), 621-637.

Preston, S. D., & De Waal, F. B. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, *25*(1), 1-20.

Qureshi, A. W. (2009). *The cognitive bases of mental state reasoning in adults* (Doctoral dissertation, University of Birmingham).

Qureshi, A.W., Apperly, I. A., & Samson, D. (2010). Executive function is necessary for perspective selection, not Level-1 visual perspective calculation: Evidence from a dual-task study of adults. *Cognition*, *117*, 230–236.

Qureshi, A.W., & Monk, R. L. (2018). Executive function underlies both perspective selection and calculation in Level-1 visual perspective taking. *Psychonomic bulletin & review*, *25*(4), 1526-1534.

Rabbitt, P., Lunn, M., Pendleton, N., Horan, M., Scott, M., Thacker, N., Lowe, C., & Jackson, A. (2007a). White matter lesions account for all age-related declines in speed but not in intelligence. *Neuropsychology*, *21*(3), 363.

Rabbitt, P., Mogapi, O., Scott, M., Thacker, N., Lowe, C., Horan, M., Pendleton, N., Jackson, A., & Lunn, D. (2007b). Effects of global atrophy, white matter lesions, and cerebral blood flow on age-related changes in speed, memory, intelligence, vocabulary, and frontal function. *Neuropsychology*, *21*(6), 684.

Rakoczy, H., Harder-Kasten, A., & Sturm, L. (2012). The decline of theory of mind in old age is (partly) mediated by developmental changes in domain-general abilities. *British Journal of Psychology*, *103*(1), 58-72.

Razza, R. A., & Blair, C. (2009). Associations among false belief understanding, executive function, and social competence: A longitudinal analysis. *Journal of Applied Developmental Psychology*, 30, 332 – 343.

Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., & Engle, R. W. (2013). No evidence of intelligence improvement after working memory training: A randomized, placebo-controlled study. *Journal of Experimental Psychology General*, 142, 359–379.

Reed, A. E., Chan, L., & Mikels, J. A. (2014). Meta-analysis of the age-related positivity effect: age differences in preferences for positive over negative information. *Psychology and aging*, *29*(1), 1.

Reuter-Lorenz, P. A., & Cappell, K. A. (2008). Neurocognitive aging and the compensation hypothesis. *Current directions in psychological science*, *17*(3), 177-182.

Reuter-Lorenz, P., & Lustig, C. (2016, 12 28). Working memory and executive functions in the aging brain. In Cognitive neuroscience of aging: Linking cognitive and cerebral aging. Oxford: Oxford University Press.

Reznick, J. S., Morrow, J. D., Goldman, B. D., & Snyder, J. (2004). The onset of working memory in infants. *Infancy*, *6*(1), 145-154.

Rhodes, M. G., & Anastasi, J. S. (2012). The own-age bias in face recognition: a meta-analytic and theoretical review. *Psychological bulletin*, *138*(1), 146.

Richmond, L. L., Morrison, A. B., Chein, J. M., & Olson, I. R. (2011). Working memory training and transfer in older adults. *Psychology and aging*, *26*(4), 813.

Richter, D., Dietzel, C., & Kunzmann, U. (2011). Age differences in emotion recognition: The task matters. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *66*(1), 48-55.

Richter, D., & Kunzmann, U. (2011). Age differences in three facets of empathy: Performance-based evidence. *Psychology and aging*, *26*(1), 60.

Risko, E. F., Laidlaw, K. E., Freeth, M., Foulsham, T., & Kingstone, A. (2012). Social attention with real versus reel stimuli: toward an empirical approach to concerns about ecological validity. *Frontiers in human neuroscience*, *6*, 143.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.*, *27*, 169-192.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2002). Motor and cognitive functions of the ventral premotor cortex. *Current opinion in neurobiology*, *12*(2), 149-154.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2009). The mirror neuron system: A motor-based mechanism for action and intention understanding.

Robinson, S., Goddard, L., Dritschel, B., Wisley, M., & Howlin, P. (2009). Executive functions in children with autism spectrum disorders. *Brain and Cognition*, *71*(3), 362-368.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, 124, 207–231.

Romine, C. B., & Reynolds, C. R. (2005). A model of the development of frontal lobe functioning: Findings from a meta-analysis. *Applied neuropsychology*, *12*(4), 190-201.

Rubia, K., Overmeyer, S., Taylor, E., Brammer, M., Williams, S. C., Simmons, A., Andrew, C., & Bullmore, E.T. (2000). Functional frontalisation with age: Mapping neurodevelopmental trajectories with fMRI. *Neuroscience and Biobehavioral Reviews*, *24*, 13–19.

Rubia, K., Smith, A. B., Woolley, J., Nosarti, C., Heyman, I., Taylor, E., & Brammer, M. (2006). Progressive increase of frontostriatal brain activation from childhood to adulthood during event-related tasks of cognitive control. *Human brain mapping*, *27*(12), 973-993.

Rubio-Fernández, P., & Geurts, B. (2013). How to pass the false-belief task before your fourth birthday. *Psychological Science*, *24*(1), 27-33.

Ruffman, T., Henry, J. D., Livingstone, V., & Phillips, L. H. (2008). A meta-analytic review of emotion recognition and aging: Implications for neuropsychological models of aging. *Neuroscience & Biobehavioral Reviews*, *32*(4), 863-881.

Russell, J. (1996). Agency: Its role in mental development. Hove, UK: Lawrence Erlbaum.

Russell, J. (1997). How executive disorders can bring about an adequate theory of mind. In J. Russell (Ed.), Autism as an executive disorder (pp. 256–304). Oxford, UK: Oxford University Press.

Sala, G., & Gobet, F. (2017). Does far transfer exist? Negative evidence from chess, music, and working memory training. *Current directions in psychological science*, *26*(6), 515-520.

Salminen, T., Strobach, T., & Schubert, T. (2012). On the impacts of working memory training on executive functioning. *Frontiers in Human Neuroscience*, *6*, 166.

Salthouse, T. A. (1991). Mediation of adult age differences in cognition by reductions in working memory and speed of processing. *Psychological Science*, *2*(3), 179-183.

Salthouse, T. A. (2009). When does age-related cognitive decline begin? *Neurobiology of Aging*, *30*(4), 507-514.

Salthouse, T. A. (2016). *Theoretical perspectives on cognitive aging*. Psychology Press.

Samson, D., Apperly, I. a, Braithwaite, J. J., Andrews, B. J., & Bodley Scott, S. E. (2010). Seeing it their way: evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology. Human Perception and Performance*, *36*(5), 1255–66.

Sandberg, P., Rönnlund, M., Nyberg, L., & Stigsdotter Neely, A. (2014). Executive process training in young and old adults. *Aging, Neuropsychology, and Cognition*, *21*(5), 577-605.

Sandoz, M., Démonet, J. F., & Fossard, M. (2014). Theory of mind and cognitive processes in aging and Alzheimer type dementia: a systematic review. *Aging & mental health*, *18*(7), 815-827.

Santiesteban, I., Kaur, S., Bird, G., & Catmur, C. (2017). Attentional processes, not implicit mentalizing, mediate performance in a perspective-taking task: Evidence from stimulation of the temporoparietal junction. *NeuroImage*, *155*, 305-311.

Santiesteban, I., Shah, P., White, S., Bird, G., & Heyes, C. (2015). Mentalizing or submentalizing in a communication task? Evidence from autism and a camera control. *Psychonomic bulletin & review*, *22*(3), 844-849.

Santiesteban, I., White, S., Cook, J., Gilbert, S. J., Heyes, C., & Bird, G. (2012). Training social cognition: from imitation to theory of mind. *Cognition*, *122*(2), 228-235.

Savitsky, K., Keysar, B., Epley, N., Carter, T., & Swanson, A. (2011). The closeness-communication bias: Increased egocentrism among friends versus strangers. *Journal of Experimental Social Psychology*, *47*(1), 269-273.

Scherf, K. S., Sweeney, J. A., & Luna, B. (2006). Brain basis of developmental change in visuospatial working memory. *Journal of cognitive neuroscience*, *18*(7), 1045-1058.

Schilbach, L. (2010). A second-person approach to other minds. *Nature Reviews Neuroscience*, *11*(6), 449-449.

Schmitt, L. M., White, S. P., Cook, E. H., Sweeney, J. A., & Mosconi, M. W. (2018). Cognitive mechanisms of inhibitory control deficits in autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, *59*(5), 586-595.

Seeman, T. E., Lusignolo, T. M., Albert, M., & Berkman, L. (2013). Social relationships, social support, and patterns of cognitive aging in healthy, high-functioning older adults: MacArthur studies of successful aging. *Health psychology*, *20*(4), 243.

Sellaro, R., Nitsche, M. A., & Colzato, L. S. (2016). The stimulated social brain: effects of transcranial direct current stimulation on social cognition. *Annals of the New York Academy of Sciences*, *1369*(1), 218-239.

Senju, A., Southgate, V., Snape, C., Leonard, M., & Csibra, G. (2011). Do 18-month-olds really attribute mental states to others? A critical test. *Psychological Science*, *22*(7), 878-880.

Shamay-Tsoory, S. G., Tibi-Elhanany, Y., & Aharon-Peretz, J. (2006). The ventromedial prefrontal cortex is involved in understanding affective but not cognitive theory of mind stories. *Social neuroscience*, *1*(3-4), 149-166.

Simons, D. J., Boot, W. R., Charness, N., Gathercole, S. E., Chabris, C. F., Hambrick, D. Z., & Stine-Morrow, E. A. L. (2016). Do “Brain-Training” Programs Work? *Psychological Science in the Public Interest*, *17*(3), 103–186.

Simpson, A. J., & Todd, A. R. (2017). Intergroup visual perspective-taking: Shared group membership impairs self-perspective inhibition but may facilitate perspective calculation. *Cognition*, *166*, 371-381.

Singer, T., & Engert, V. (2019). It matters what you practice: Differential training effects on subjective experience, behavior, brain and body in the ReSource Project. *Current opinion in psychology*, *28*, 151-158.

Singer, T., Seymour, B., O'doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, *303*(5661), 1157-1162.

Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill* (No. 9). Harvard University Press.

Singh-Manoux, A., Kivimaki, M., Glymour, M. M., Elbaz, A., Berr, C., Ebmeier, K. P., Ferrie, J.E., & Dugravot, A. (2012). Timing of onset of cognitive decline: results from Whitehall II prospective cohort study. *Bmj*, *344*, d7622.

Slaughter, V., & Gopnik, A. (1996). Conceptual coherence in the child's theory of mind: Training children to understand belief. *Child development*, *67*(6), 2967-2988.

Slessor, G., Laird, G., Phillips, L. H., Bull, R., & Filippou, D. (2010). Age-related differences in gaze following: Does the age of the face matter?. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *65*(5), 53

Slessor, G., Phillips, L. H., Ruffman, T., Bailey, P. E., & Insch, P. (2014). Exploring own-age biases in deception detection. *Cognition & emotion*, *28*(3), 493-506. 6-541.

Sodian, B., Thoermer, C., & Metz, U. (2007). Now I see it but you don't: 14-month-olds can represent another person's visual perspective. *Developmental science*, *10*(2), 199-204.

Solesio-Jofre, E., Lorenzo-López, L., Gutiérrez, R., López-Frutos, J. M., Ruiz-Vargas, J. M., & Maestú, F. (2012). Age-related effects in working memory recognition modulated by retroactive interference. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, *67*(6), 565-572.

Soveri, A., Antfolk, J., Karlsson, L., Salo, B., & Laine, M. (2017). Working memory training revisited: A multi-level meta-analysis of n-back training studies. *Psychonomic bulletin & review*, *24*(4), 1077-1096.

Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human perception and performance*, *22*(2), 461.

Spooner, J. W., Sakala, S. M., & Baloh, R. W. (1980). Effect of aging on eye tracking. *Archives of Neurology*, *37*(9), 575-576.

Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. *Developmental review*, *28*(1), 78-106.

Stine-Morrow, E. A., & Basak, C. (2011). Cognitive interventions. In *Handbook of the psychology of aging* (pp. 153-171). Academic Press.

Stopford, C. L., Thompson, J. C., Neary, D., Richardson, A. M., & Snowden, J. S. (2012). Working memory, attention, and executive function in Alzheimer’s disease and frontotemporal dementia. *Cortex*, *48*(4), 429-446.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of experimental psychology*, *18*(6), 643.

Sucksmith, E., Allison, C., Baron-Cohen, S., Chakrabarti, B., & Hoekstra, R. A. (2013). Empathy and emotion recognition in people with autism, first-degree relatives, and controls. *Neuropsychologia*, *51*(1), 98-105.

Sullivan, S., & Ruffman, T. (2004a). Social understanding: How does it fare with advancing years? *British Journal of Psychology*, *95*(1), 1-18.

Sullivan, S., & Ruffman, T. (2004b). Emotion recognition deficits in the elderly. *International Journal of Neuroscience*, *114*(3), 403-432.

Sullivan, S., Campbell, A., Hutton, S. B., & Ruffman, T. (2017). What's good for the goose is not good for the gander: Age and gender differences in scanning emotion faces. *The Journals of Gerontology: Series B*, *72*(3), 441-447.

Sun, B., Luo, Z., Zhang, W., Li, W., Li, X. (2018). Age-related differences in affective and cognitive empathy: self-report and performance-based evidence. *Aging, Neuropsychology, and Cognition*, *25*(5), 655-672

Surtees, A. D., & Apperly, I. A. (2012). Egocentrism and automatic perspective taking in children and adults. *Child Development*, *83*(2), 452-460.

Surtees, A., Apperly, I., & Samson, D. (2013a). Similarities and differences in visual and spatial perspective-taking processes. *Cognition*, *129*(2), 426-438.

Surtees, A. D. R., Apperly, I. A., & Samson, D. (2013b). The use of embodied self-rotation for visual and spatial perspective-taking. *Frontiers in human neuroscience*, *7*, 698.

Surtees, A., Samson, D., & Apperly, I. (2016). Unintentional perspective-taking calculates whether something is seen, but not how it is seen. *Cognition*, *148*, 97-105.

Symeonidou, I., Dumontheil, I., Chow, W. Y., & Breheny, R. (2016). Development of online use of theory of mind during adolescence: An eye-tracking study. *Journal of experimental child psychology*, *149*, 81-97.

Symeonidou, I., Ferguson, H. J., Dumontheil, I., & Breheny, R. (2020). Author accepted manuscript: Adolescents are delayed at inferring complex social intentions in others, but not basic (false) beliefs: An eye movement investigation. *Quarterly Journal of Experimental Psychology*, 1747021820920213.

Sze, J. A., Gyurak, A., Goodkind, M. S., & Levenson, R. W. (2012). Greater emotional empathy and prosocial behavior in late life. *Emotion, 12*(5), 1129–1140.

Taylor, S. J., Barker, L. A., Heavey, L., & McHale, S. (2013). The typical developmental trajectory of social and executive functions in late adolescence and early adulthood. *Developmental Psychology*, *49*(7), 1253.

Taylor, S. J., Barker, L. A., Heavey, L., & McHale, S. (2015). The longitudinal development of social and executive functions in late adolescence and early adulthood. *Frontiers in Behavioral Neuroscience*, *9*, 252.

Teding van Berkhout, E., & Malouff, J. M. (2016). The efficacy of empathy training: A meta-analysis of randomized controlled trials. *Journal of counseling psychology, 63*(1), 32.

Thorell, L. B., Lindqvist, S., Bergman Nutley, S., Bohlin, G., & Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Developmental Science, 12*: 106-113.

Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions. II. The estimation of magnitudes. *Psychological Review*, *8*(4), 384.

Todd, A. R., Cameron, C. D., & Simpson, A. J. (2017). Dissociating processes underlying level-1 visual perspective taking in adults. *Cognition*, *159*, 97-101.

Todd, A. R., Hanko, K., Galinsky, A. D., & Mussweiler, T. (2011). When focusing on differences leads to similar perspectives. *Psychological Science*, *22*(1), 134-141.

Towse, J. N., Hitch, G. J., & Hutton, U. (2000). On the interpretation of working memory span in adults. *Memory & Cognition*, *28*(3), 341-348.

Towse, J. N., Hitch, G. J., & Skeates, S. (1999). Developmental sensitivity to temporal grouping effects in short-term memory. *International Journal of Behavioral Development*, *23*(2), 391-411.

Trautwein, F. M., Kanske, P., Böckler, A., & Singer, T. (2020). Differential benefits of mental training types for attention, compassion, and theory of mind. *Cognition*, *194*, 104039.

Uekermann, J., Abdel-Hamid, M., Lehmkamper, C., Vollmoeller, W., Daum, I., (2008a). Perception of affective prosody in major depression: a link to executive functions? *Journal of the International Neuropsychological Society, 14*, 552–561.

Vabalas, A., & Freeth, M. (2016). Brief report: Patterns of eye movements in face to face conversation are associated with autistic traits: Evidence from a student sample. *Journal of autism and developmental disorders*, *46*(1), 305-314.

van den Bos, W., van Dijk, E., Westenberg, M., Rombouts, S. A., & Crone, E. A. (2011). Changing brains, changing perspectives: the neurocognitive development of reciprocity. *Psychological science*, *22*(1), 60-70.

van den Heuvel, O. A., Van Gorsel, H. C., Veltman, D. J., & Van Der Werf, Y. D. (2013). Impairment of executive performance after transcranial magnetic modulation of the left dorsal frontal-striatal circuit. *Human brain mapping*, *34*(2), 347-355.

van den Wildenberg, W. P., & van der Molen, M. W. (2004). Developmental trends in simple and selective inhibition of compatible and incompatible responses. *Journal of experimental child psychology*, *87*(3), 201-220.

Verbaan, D., Marinus, J., Visser, M., van Rooden, S. M., Stiggelbout, A. M., & van Hilten, J. J. (2007). Patient-reported autonomic symptoms in Parkinson disease. *Neurology*, *69*(4), 333-341.

Verhaeghen, P. (2011). Aging and executive control: Reports of a demise greatly exaggerated. *Current Directions in Psychological Science, 20,* 174–180.

Vetter, N. C., Leipold, K., Kliegel, M., Phillips, L. H., & Altgassen, M. (2013). Ongoing development of social cognition in adolescence. *Child Neuropsychology*, *19*(6), 615-629.

Vicaria, I. M., Bernieri, F. J., & Isaacowitz, D. M. (2015). Perceptions of rapport across the life span: Gaze patterns and judgment accuracy. *Psychology and Aging*, *30*(2), 396.

Von Bastian, C. C., & Eschen, A. (2016). Does working memory training have to be adaptive? *Psychological research*, *80*(2), 181-194.

Voss, M. W., Prakash, R. S., Erickson, K. I., Basak, C., Chaddock, L., Kim, J. S., Alves H., Heo S., Szabo, A., White, S.M., & Wójcicki, T. R. (2010). Plasticity of brain networks in a randomized intervention trial of exercise training in older adults. *Frontiers in aging neuroscience*, *2*, 32.

Wade, M., Prime, H., Jenkins, J. M., Yeates, K. O., Williams, T., & Lee, K. (2018). On the relation between theory of mind and executive functioning: A developmental cognitive neuroscience perspective. *Psychonomic Bulletin & Review*, *25*(6), 2119-2140.

Wasylyshyn, C., Verhaeghen, P., & Sliwinski, M. J. (2011). Aging and task switching: A meta-analysis. *Psychology and Aging*, *26*(1), 15–20.

Watkins, L. H., Sahakian, B. J., Robertson, M. M., Veale, D. M., Rogers, R. D., Pickard, K. M., Aitken, M.R., & Robbins, T. W. (2005). Executive function in

Tourette's syndrome and obsessive–compulsive disorder. *Psychological medicine*, *35*(4), 571-582.

Wechsler, D. (1999). Wechsler Abbreviated Scale of Intelligence. San Antonio, TX: The Psychological Corporation.

Weil, L. G., Fleming, S. M., Dumontheil, I., Kilford, E. J., Weil, R. S., Rees, G.,Dolan R.J., & Blakemore, S. J. (2013). The development of metacognitive ability in adolescence. *Consciousness and cognition*, *22*(1), 264-271.

Wellman, H. M. (2002). *Understanding the psychological world: Developing a theory of mind*. In U. Goswami (Ed.), Blackwell handbook of childhood cognitive development (pp. 167–187). Malden: Blackwell.

Wellman, H., Cross, D., & Watson, J. (2001). Meta-analysis of theory of mind development: The truth about false belief. *Child Development, 72*, 655 – 684.

Wellman, H. M., & Woolley, J. D. (1990). From simple desires to ordinary beliefs: The early development of everyday psychology. *Cognition*, *35*, 245–275.

Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: a meta-analytic review. *Biological psychiatry*, *57*(11), 1336-1346.

Williams, B. T., Gray, K. M., & Tonge, B. J. (2012). Teaching emotion recognition skills to young children with autism: a randomised controlled trial of an emotion training programme. *Journal of Child Psychology and Psychiatry*, *53*(12), 1268-1276.

Wilkinson, A.J., & Yang, L. (2012). Plasticity of inhibition in older adults: Retest practice and transfer effects. *Psychology and Aging, 27*, 606 – 615.

Wilson, M. (1988). MRC psycholinguistic database: Machine-usable dictionary, version 2.00. *Behavior research methods, instruments, & computers*, *20*(1), 6-10.

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

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69,* 1–34

Woodworth, R. S., & Thorndike, E. L. (1901). The influence of improvement in one mental function upon the efficiency of other functions. *Psychological review*, *8*(3), 247.

Wright, J. C., & Mischel, W. (1987). A conditional approach to dispositional constructs: the local predictability of social behavior. *Journal of personality and social psychology*, *53*(6), 1159.

Yan, Z., Hong, S., Liu, F., & Su, Y. (2020). A meta-analysis of the relationship between empathy and executive function. *PsyCh Journal*, *9*(1), 34-43.

Yang, C. Y., Decety, J., Lee, S., Chen, C., & Cheng, Y. (2009). Gender differences in the mu rhythm during empathy for pain: an electroencephalographic study. *Brain research*, *1251*, 176-184.

Yıldırım, E., Soncu Büyükişcan, E., & Gürvit, H. (2020). Affective theory of mind in human aging: is there any relation with executive functioning?. *Aging, Neuropsychology, and Cognition, 27*(2), 207-219.

Zanto, T.P., Hennigan, K., Ostberg, M., Clapp, W.C., & Gazzaley, A. (2010). Predictive knowledge of stimulus relevance does not influence top-down suppression of irrelevant information in older adults. *Cortex 46*:564–74.

Zappasodi, F., Marzetti, L., Olejarczyk, E., Tecchio, F., & Pizzella, V. (2015). Age-related changes in electroencephalographic signal complexity. *PloS one*, *10*(11), e0141995

Ze, O., Thoma, P., & Suchan, B. (2014). Cognitive and affective empathy in younger and older individuals. *Aging & mental health*, *18*(7), 929-935.

Zhang, X., Lecce, S., Ceccato, I., Cavallini, E., Zhang, L., & Chen, T. (2018). Plasticity in older adults' theory of mind performance: The impact of motivation. *Aging & Mental Health*, *22*(12), 1592-1599.

Zhao, X., Chen, L., & Maes, J. H. (2018). Training and transfer effects of response inhibition training in children and adults. *Developmental Science*, *21*(1), e12511.

Zhu, X., Yin, S., Lang, M., He, R., & Li, J. (2016). The more the better? A meta-analysis on effects of combined cognitive and physical intervention on cognition in healthy older adults. *Ageing Research Reviews, 31*, 67-79.

Zinke K., Einert M., Pfennig L., & Kliegel M. (2012). Plasticity of executive control through task switching training in adolescents. *Front. Hum. Neurosci.* 6:41

Zinke, K., Zeintl, M., Rose, N. S., Putzmann, J., Pydde, A., & Kliegel, M. (2014). Working memory training and transfer in older adults: effects of age, baseline performance, and training gains. *Developmental psychology*, *50*(1), 304.

Zobel, I., Werden, D., Linster, H., Dykierek, P., Drieling, T., Berger, M., & Schramm, E. (2010). Theory of mind deficits in chronically depressed patients. *Depression and Anxiety* 2*7*, 821–828.

Appendix A

Physical Pain Stimuli used in the empathy for others' pain task.




















No Physical Pain Stimuli used in the empathy for others’ pain task.



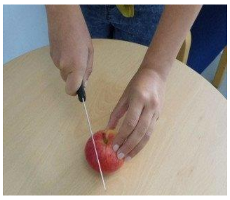








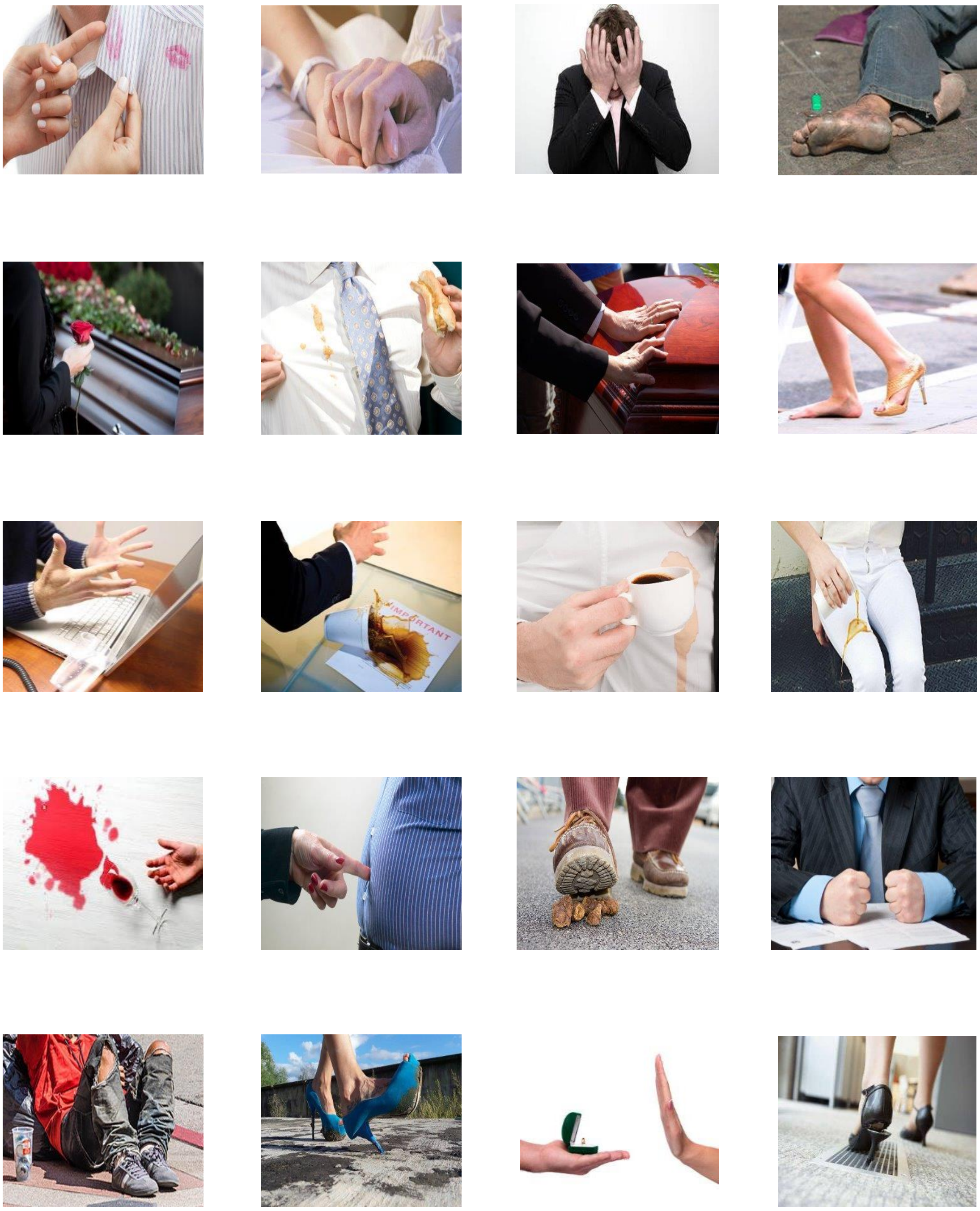









Social Pain Stimuli used in the empathy for others' pain task.



No Social Pain Stimuli used in the empathy for others’ pain task.




















Appendix B

Full set of experimental questions used in the face to face conversation task. Note that on set was used at pre-test and one was used at post-test.

Set A

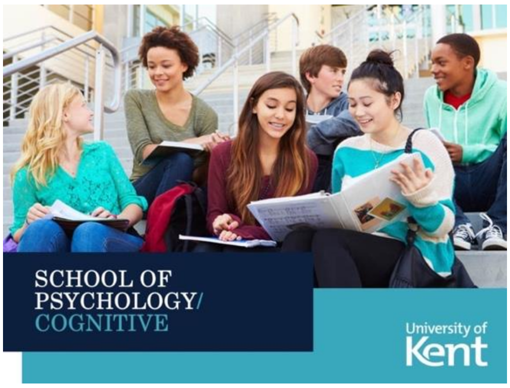
1. Tell me some things you like about living in Kent and some things you dislike about living in Kent.
2. Tell me about some things that you did last weekend and some things that you plan to do next weekend.
3. Describe a few things you consider to be typically English and a few things you consider to be typically American.
4. Tell me about some things you do in your spare time; then pick one sport or activity of your choice and either describe some of the rules or tell me how you would go about doing that sport or activity.

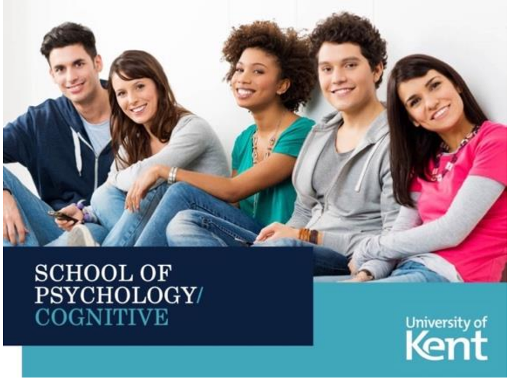
Set B

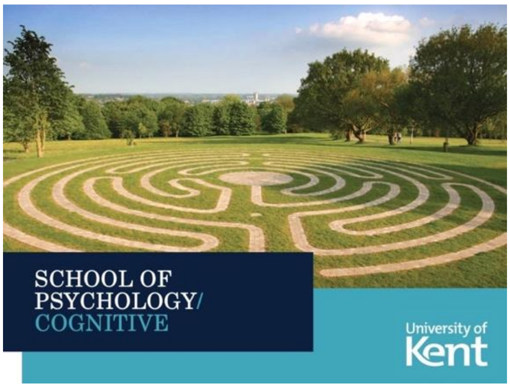
1. Tell me some things you like about Christmas and some things you dislike about Christmas.
2. Tell me about some things that you did last weekend and some things that you plan to do next weekend.
3. Describe a few foods that you consider to be typically Scottish and a few foods that you consider to be typically Italian.
4. Tell me about a TV show or movie that you’ve watched recently; then pick one and describe either a character or a recent storyline.

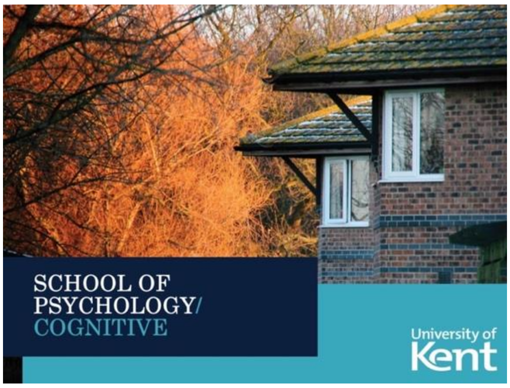
Appendix C

Full set of posters used as background stimuli in the face-to-face conversation task. Two posters were created for each of the three conditions (i.e. social scenes with averted gaze or shared gaze, and one non-social scene)- one set was used at pre-test and one was used at post-test.









Appendix D

Statistical effects for accuracy in the visuo-spatial perspective-taking task. Asterisks show significance of effects, where * $p < .05$; *** $p < .001$.

	df	*F*	*P*	η_p^2
Age Group	(2, 279)	20.41	< .001***	.13
Content	(1, 279)	14.38	< .001***	.05
Angle	(3, 837)	71.63	< .001***	.20
Level	(1, 279)	8.90	.003**	.03
Age Group x Content	(2, 279)	0.05	.95	< .01
Age Group x Angle	(6, 837)	6.50	< .001***	.04
Age Group x Level	(2, 279)	0.11	.89	< .01
Content x Angle	(3, 837)	5.84	< .001***	.02
Content x Level	(1, 279)	36.30	< .001***	.12
Angle x Level	(3, 837)	29.23	< .001***	.10
Age Group x Content x Angle	(6, 837)	1.45	.20	.01
Age Group x Content x Level	(2, 279)	0.91	.40	< .01
Age Group x Angle x Level	(6, 837)	3.65	.001***	.03
Content x Angle x Level	(3, 837)	1.0	.39	< .01
Age Group x Content x Angle x Level	(6, 837)	1.60	.14	.01